(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 April 2001 (05.04.2001)

PCT

(10) International Publication Number
WO 01/23410 A2

(51) International Patent Classification[7]: C07K

(21) International Application Number: PCT/US00/26556

(22) International Filing Date: 27 September 2000 (27.09.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/156,489 28 September 1999 (28.09.1999) US

(71) Applicant *(for all designated States except US)*: **GENAISSANCE PHARMACEUTICALS, INC.** [US/US]; Fivie Science Park, New Haven, CT 06511 (US).

(72) **Inventors; and**
(75) **Inventors/Applicants** *(for US only)*: **CHEW, Anne** [US/US]; 361 Fountain Street, New Haven, CT 06511 (US). **DENTON, R., Rex** [US/US]; 129 Hunters Trail, Madison, CT 06443 (US). **NANDABALAN, Krishnan** [IN/US]; 228 Village Pond Road, Guilford, CT 06437 (US). **STEPHENS, Joel, Claiborne** [US/US]; 46 Crabapple Lane, Guilford, CT 06437 (US).

(74) **Agents: SHTIVELBAND, Inna** et al.; Genaissance Pharmaceuticals, Inc., Five Science Park, New Haven, CT 06511 (US).

(81) **Designated States** *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/23410 A2

(54) **Title:** DRUG TARGET ISOGENES: POLYMORPHISMS IN THE INTERLEUKIN 13 GENE

(57) **Abstract:** Polynucleotides comprising one or more of 14 novel single nucleotide polymorphisms in the human interleukin 13 (IL13) gene are described. Compositions and methods for detecting one or more of these polymorphisms are also disclosed. In addition, various genotypes and haplotypes for IL13 gene that exist in the population are described.

DRUG TARGET ISOGENES:
POLYMORPHISMS IN THE INTERLEUKIN 13 GENE

RELATED APPLICATIONS

This application is a continuation-in-part of, and claims priority to, U.S. Provisional Application Serial No. 60/156,489 filed September 28, 1999.

FIELD OF THE INVENTION

This invention relates to variation in genes that encode pharmaceutically important proteins. In particular, this invention provides genetic variants of the human Interleukin 13 (IL13) gene and methods for identifying which variant(s) of this gene is/are possessed by an individual.

BACKGROUND OF THE INVENTION

Current methods for identifying pharmaceuticals to treat disease often start by identifying, cloning, and expressing an important target protein related to the disease. A determination of whether an agonist or antagonist is needed to produce an effect that may benefit a patient with the disease is then made. Then, vast numbers of compounds are screened against the target protein to find new potential drugs. The desired outcome of this process is a drug that is specific for the target, thereby reducing the incidence of the undesired side effects usually caused by a compound's activity at non-intended targets.

What this approach fails to consider, however, is that natural variability exists in any and every population with respect to a particular protein. A target protein currently used to screen drugs typically is expressed by a gene cloned from an individual who was arbitrarily selected. However, the nucleotide sequence of a particular gene may vary tremendously among individuals. Subtle alteration(s) in the primary nucleotide sequence of a gene encoding a target protein may be manifested as significant variation in expression of or in the structure and/or function of the protein. Such alterations may explain the relatively high degree of uncertainty inherent in treatment of individuals with drugs whose design is based upon a single representative example of the target. For example, it is well-established that some classes of drugs frequently have lower efficacy in some individuals than others, which means such individuals and their physicians must weigh the possible benefit of a larger dosage against a greater risk of side effects. In addition, variable information on the biological function or effects of a particular protein may be due to different scientists unknowingly studying different isoforms of the gene encoding the protein. Thus, information on the type and frequency of genomic variation that exists for pharmaceutically important proteins would be useful.

The organization of single nucleotide variations (polymorphisms) in the primary sequence of a gene into one of the limited number of combinations that exist as units of inheritance is termed a haplotype. Each haplotype therefore contains significantly more information than individual

unorganized polymorphisms. Haplotypes provide an accurate measurement of the genomic variation in the two chromosomes of an individual.

It is well-established that many diseases are associated with specific variations in gene sequences. However while there are examples in which individual polymorphisms act as genetic markers for a particular phenotype, in other cases an individual polymorphism may be found in a variety of genomic backgrounds and therefore shows no definitive coupling between the polymorphism and the causative site for the phenotype (Clark AG et al. 1998 *Am J Hum Genet* **63**:595-612; Ulbrecht M et al. 2000 *Am J Respir Crit Care Med* **161**: 469-74). In addition, the marker may be predictive in some populations, but not in other populations (Clark AG et al. 1998 *supra*). In these instances, a haplotype will provide a superior genetic marker for the phenotype (Clark AG et al. 1998 *supra*; Ulbrecht M et al. 2000, *supra*; Ruaño G & Stephens JC *Gen Eng News* **19** (21), December 1999).

Analysis of the association between each observed haplotype and a particular phenotype permits ranking of each haplotype by its statistical power of prediction for the phenotype. Haplotypes found to be strongly associated with the phenotype can then have that positive association confirmed by alternative methods to minimize false positives. For a gene suspected to be associated with a particular phenotype, if no observed haplotypes for that gene show association with the phenotype of interest, then it may be inferred that variation in the gene has little, if any, involvement with that phenotype (Ruaño & Stephens 1999, *supra*). Thus, information on the observed haplotypes and their frequency of occurrence in various population groups will be useful in a variety of research and clinical applications.

One possible drug target for the treatment of cancers, inflammatory and immune disorders is the Interleukin 13 (IL13) gene or its encoded product. IL13 is a pro-inflammatory cytokine expressed by activated human T lymphocytes (Minty et al., *Nature* 362:248-250, 1993; McKenzie et al., *Proc. Natl. Acad. Sci. USA* 90:3735-3739, 1993). IL-13 shares some biological activities with IL4 and these two T-cell cytokines may also share a common signaling pathway through the IL-4 receptor (Punnonen et al., *Proc. Natl. Acad. Sci.* 90:3730-3734, 1993). Several recent reports have suggested a role for IL13 in the pathogenesis of asthma. For example, IL-13 was found to be necessary and sufficient for the expression of allergic asthma in mice (Wills-Karp et al., *Science* 282:2258-2261, 1998; Grunig et al., *Science* 282:2261-2263, 1998) and targeted pulmonary expression of IL-13 in transgenic mice produced responses similar to those observed in asthma (Zhu et al., *J. Clin. Invest.* 103:779-788, 1999). In addition, IL-13 expression is elevated in the airways of both allergic and nonallergic asthma patients (Walker et al., *Am. Rev. Respir. Dis.* 146:109, 1992; Humbert et al, *J. Allergy Clin. Immunol.* 99:657, 1997; Huang, S.K., *J. Immunol.* 155:2688, 1995). Finally, human asthma has been linked genetically to the region of chromosome 5q that contains both the IL-13 and IL-4 genes (Marsh et al., *Science* 264:1152, 1996).

IL13 has also been implicated in the pathogenesis of diseases characterized by fibrosis.

Studies on normal (NF) and keloid (KF) fibroblasts have shown that IL13 stimulates collagen production in both NF and KF cells by the expression of genes for procollagen 1 alpha 1. In NF, IL13 inhibited matrix metalloproteinase-3 (MMP-3) production and enhanced tissue inhibitor of metalloproteinase (TIMP)-1 generation, whereas in KF, IL13 enhanced MMP-3 production but the levels of TIMP-1 remained unaffected. These data show the importance of IL13 in collagen homeostasis and the differential regulation of its effectors in fibroblasts (Oriente et al., *J. Pharmacol. Exp.* Ther 2000 Mar;292(3):988-94).

Kanai et al., (*Br. J. Cancer* 2000 May;82(10):1717-23) studied the role of IL13 in colon cancer and found that IL13 inhibited colon cancer cell-cell adhesion by down-regulating the expression of E-cadherin and carcinoembryonic antigen (CEA) molecules.

In the vast majority of patients with high-grade gliomas (HGG), mutated IL13-based cytotoxins, directed specifically towards glioma-associated sites, are known to be the most active anti-glioma agents (Debinski W and Gibo DM *Mol Med* 2000 May;6(5):440-9). A mutagenized IL13, known as hIL13.E13K (glutamic acid at position 13 changed to lysine), was fused to Pseudomonas exotoxin A. This mutagenized IL13 was less active and less toxic on normal cells than gliomas and therefore served as an antitumor agent (Debinski et al., *Nat. Biotechnol* 1998 may;16(5):449-53).

The Interleukin 13 gene is located on chromosome 5q31 and contains 4 exons that encode a 131 amino acid protein. Reference sequences for the IL13 gene (SEQ ID NO:1), coding sequence, and protein are shown in Figures 1, 2 and 3, respectively.

One single nucleotide polymorphism of a guanine or adenine at nucleotide position 4165 in Figure 1, reported in the NCBI SNP Database (Ref SNP#20541), released on December 21, 1999 was originally reported in U.S Provisional Application Serial No._60/156,489 filed September 28, 1999.

This polymorphism results in a variation of arginine or glutamine at amino acid position 129 in Figure 3. Heinzmann et al. (*Hum. Mol. Genet.* 2000 Mar 1;9(4):549-59) have shown that this Arg129Gln variant affects the interaction of IL13 with its receptor and was found to be associated with asthma in their study population from Britain and Japan. Because of the potential for polymorphisms in the IL13 gene to affect the expression and function of the encoded protein, it would be useful to determine whether additional polymorphisms exist in the IL13 gene, as well as how such polymorphisms are combined in different copies of the gene. Such information would be useful for studying the biological function of IL13 as well as in identifying drugs targeting this protein for the treatment of disorders related to its abnormal expression or function.

SUMMARY OF THE INVENTION

Accordingly, the inventors herein have discovered 14 novel polymorphic sites in the IL13 gene. These polymorphic sites (PS) correspond to the following nucleotide positions in Figure 1: 1100 (PS1), 1408 (PS2), 1477 (PS3), 1842 (PS4), 2080 (PS5), 2233 (PS6), 2385 (PS7), 3248 (PS8), 3280 (PS9), 3807 (PS10), 4014 (PS11), 4033 (PS12), 4044 (PS13) and 4165 (PS14). The

polymorphisms at these sites are cytosine or thymine at PS1, adenine or guanine at PS2, adenine or guanine at PS3, cytosine or thymine at PS4, thymine or cytosine at PS5, cytosine or adenine at PS6, cytosine or thymine at PS7, guanine or adenine at PS8, cytosine or thymine at PS9, cytosine or thymine at PS10, guanine or adenine at PS11, cytosine or thymine at PS12, cytosine or thymine at PS13 and guanine or adenine at PS14. In addition, the inventors have determined the identity of the alternative nucleotides present at these sites in a human reference population of 79 unrelated individuals self-identified as belonging to one of four major population groups: African descent, Asian, Caucasian and Hispanic/Latino. It is believed that IL13-encoding polynucleotides containing one or more of the novel polymorphic sites reported herein will be useful in studying the expression and biological function of IL13, as well as in developing drugs targeting this protein. In addition, information on the combinations of polymorphisms in the IL13 gene may have diagnostic and forensic applications.

Thus, in one embodiment, the invention provides an isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the IL13 gene or a fragment thereof. The reference sequence comprises SEQ ID NO:1 and the polymorphic variant comprises at least one polymorphism selected from the group consisting of thymine at PS1, guanine at PS2, guanine at PS3, thymine at PS4, cytosine at PS5, adenine at PS6, thymine at PS7, adenine at PS8, thymine at PS9, thymine at PS10, adenine at PS11, thymine at PS12, thymine at PS13 and adenine at PS14. A particularly preferred polymorphic variant is a naturally-occurring isoform (also referred to herein as an "isogene") of the IL13 gene. An IL13 isogene of the invention comprises cytosine or thymine at PS1, adenine or guanine at PS2, adenine or guanine at PS3, cytosine or thymine at PS4, thymine or cytosine at PS5, cytosine or adenine at PS6, cytosine or thymine at PS7, guanine or adenine at PS8, cytosine or thymine at PS9, cytosine or thymine at PS10, guanine or adenine at PS11, cytosine or thymine at PS12, cytosine or thymine at PS13 and guanine or adenine at PS14. The invention also provides a collection of IL13 isogenes, referred to herein as an IL13 genome anthology.

An IL13 isogene may be defined by the combination and order of these polymorphisms in the isogene, which is referred to herein as an IL13 haplotype. Thus, the invention also provides data on the number of different IL13 haplotypes found in the above four population groups. This haplotype data is useful in methods for deriving an IL13 haplotype from an individual's genotype for the IL13 gene and for determining an association between an IL13 haplotype and a particular trait.

In another embodiment, the invention provides a polynucleotide comprising a polymorphic variant of a reference sequence for an IL13 cDNA or a fragment thereof. The reference sequence comprises SEQ ID NO:2 (Fig. 2) and the polymorphic cDNA comprises at least one polymorphism selected from the group consisting of adenine at a position corresponding to nucleotide 111 and adenine at a position corresponding to nucleotide 386.

Polynucleotides complementary to these IL13 genomic and cDNA variants are also provided by the invention.

In other embodiments, the invention provides a recombinant expression vector comprising one of the polymorphic genomic variants operably linked to expression regulatory elements as well as a recombinant host cell transformed or transfected with the expression vector. The recombinant vector and host cell may be used to express IL13 for protein structure analysis and drug binding studies.

In yet another embodiment, the invention provides a polypeptide comprising a polymorphic variant of a reference amino acid sequence for the IL13 protein. The reference amino acid sequence comprises SEQ ID NO:3 (Fig. 3) and the polymorphic variant comprises glutamine at a position corresponding to amino acid position 129. A polymorphic variant of IL13 is useful in studying the effect of the variation on the biological activity of IL13 as well as studying the binding affinity of candidate drugs targeting IL13 for the treatment of cancers, inflammatory and immune disorders.

The present invention also provides antibodies that recognize and bind to the above polymorphic IL13 protein variant. Such antibodies can be utilized in a variety of diagnostic and prognostic formats and therapeutic methods.

In other embodiments, the invention provides methods, compositions, and kits for haplotyping and/or genotyping the IL13 gene in an individual. The methods involve identifying the nucleotide or nucleotide pair present at one or more polymorphic sites selected from PS1-PS14 in one or both copies of the IL13 gene from the individual. The compositions contain oligonucleotide probes and primers designed to specifically hybridize to one or more target regions containing, or that are adjacent to, a polymorphic site. The methods and compositions for establishing the genotype or haplotype of an individual at the novel polymorphic sites described herein are useful for studying the effect of the polymorphisms in the etiology of diseases affected by the expression and function of the IL13 protein, studying the efficacy of drugs targeting IL13, predicting individual susceptibility to diseases affected by the expression and function of the IL13 protein and predicting individual responsiveness to drugs targeting IL13.

In yet another embodiment, the invention provides a method for identifying an association between a genotype or haplotype and a trait. In preferred embodiments, the trait is susceptibility to a disease, severity of a disease, the staging of a disease or response to a drug. Such methods have applicability in developing diagnostic tests and therapeutic treatments for cancers, inflammatory and immune disorders.

The present invention also provides transgenic animals comprising one of the IL13 genomic polymorphic variants described herein and methods for producing such animals. The transgenic animals are useful for studying expression of the IL13 isogenes *in vivo*, for *in vivo* screening and testing of drugs targeted against IL13 protein, and for testing the efficacy of therapeutic agents and compounds for cancers, inflammatory and immune disorders in a biological system.

The present invention also provides a computer system for storing and displaying polymorphism data determined for the IL13 gene. The computer system comprises a computer processing unit; a display; and a database containing the polymorphism data. The polymorphism data

includes the polymorphisms, the genotypes and the haplotypes identified for the IL13 gene in a reference population. In a preferred embodiment, the computer system is capable of producing a display showing IL13 haplotypes organized according to their evolutionary relationships.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates a reference sequence for the IL13 gene (contiguous lines; SEQ ID NO:1), with the start and stop positions of each region of coding sequence indicated below the sequence by the numbers within the brackets and the polymorphic sites and polymorphisms identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the sequence.

Figure 2 illustrates a reference sequence for the IL13 coding sequence (contiguous lines; SEQ ID NO:2), with the polymorphic sites and polymorphisms identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the sequence.

Figure 3 illustrates a reference sequence for the IL13 protein (contiguous lines; SEQ ID NO:3), with the variant amino acids caused by the polymorphisms of Fig. 2 positioned below the polymorphic site in the sequence.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is based on the discovery of novel variants of the IL13 gene. As described in more detail below, the inventors herein discovered 14 novel polymorphic sites by characterizing the IL13 gene found in genomic DNAs isolated from an Index Repository IA that contains immortalized cell lines from one chimpanzee and 93 human individuals and Index Repository 1B that contains 70 individuals. These two repositories contain 10 individuals in common.

The human individuals in Index Repository IA included a reference population of 79 unrelated individuals self-identified as belonging to one of four major population groups: Caucasian (22 individuals), African descent (20 individuals) Asian (20 individuals) Hispanic/Latino (17 individuals). To the extent possible, the members of this reference population were organized into population subgroups by the self-identified ethnogeographic origin of their four grandparents as shown in Table 1 below.

Table 1. Population Groups in the Index Repository

Population Group	Population Subgroup	No. of Individuals
African descent		20
	Sierra Leone	1
Asian		20
	Burma	1
	China	3
	Japan	6
	Korea	1
	Philippines	5
	Vietnam	4
Caucasian		22
	British Isles	3
	British Isles/Central	4
	British Isles/Eastern	1
	Central/Eastern	1
	Eastern	3
	Central/Mediterranean	1
	Mediterranean	2
	Scandinavian	2
Hispanic/Latino		17
	Caribbean	7
	Caribbean (Spanish Descent)	2
	Central American (Spanish Descent)	1
	Mexican American	4
	South American (Spanish Descent)	3

In addition, the Index Repository contains three unrelated indigenous American Indians (one from each of North, Central and South America), one three-generation Caucasian family (from the CEPH Utah cohort) and one two-generation African-American family. Index Repository IB contains a reference population of 70 human individuals comprised of 4 three-generation families (from the CEPH Utah cohort) as well as unrelated African-American, Asian and Caucasian individuals. A total of 38 individuals in this reference population are unrelated.

Using the IL13 genotypes identified in the Index Repository and the methodology described in the Examples below, the inventors herein also determined the haplotypes found on each chromosome for most human members of this repository. The IL13 genotypes and haplotypes found in the repository include those shown in Tables 3 and 4, respectively. The polymorphism and haplotype data disclosed herein are useful for studying population diversity, anthropological lineage, the significance of diversity and lineage at the phenotypic level, paternity testing, forensic applications, and for identifying associations between the IL13 genetic variation and a trait such as level of drug response or susceptibility to disease.

In the context of this disclosure, the following terms shall be defined as follows unless otherwise indicated:

Allele - A particular form of a genetic locus, distinguished from other forms by its particular

nucleotide sequence.

Candidate Gene – A gene which is hypothesized to be responsible for a disease, condition, or the response to a treatment, or to be correlated with one of these.

Gene - A segment of DNA that contains all the information for the regulated biosynthesis of an RNA product, including promoters, exons, introns, and other untranslated regions that control expression.

Genotype – An unphased 5′ to 3′ sequence of nucleotide pair(s) found at one or more polymorphic sites in a locus on a pair of homologous chromosomes in an individual. As used herein, genotype includes a full-genotype and/or a sub-genotype as described below.

Full-genotype – The unphased 5′ to 3′ sequence of nucleotide pairs found at all known polymorphic sites in a locus on a pair of homologous chromosomes in a single individual.

Sub-genotype – The unphased 5′ to 3′ sequence of nucleotides seen at a subset of the known polymorphic sites in a locus on a pair of homologous chromosomes in a single individual.

Genotyping – A process for determining a genotype of an individual.

Haplotype – A 5′ to 3′ sequence of nucleotides found at one or more polymorphic sites in a locus on a single chromosome from a single individual. As used herein, haplotype includes a full-haplotype and/or a sub-haplotype as described below.

Full-haplotype – The 5′ to 3′ sequence of nucleotides found at all known polymorphic sites in a locus on a single chromosome from a single individual.

Sub-haplotype – The 5′ to 3′ sequence of nucleotides seen at a subset of the known polymorphic sites in a locus on a single chromosome from a single individual.

Haplotype pair – The two haplotypes found for a locus in a single individual.

Haplotyping – A process for determining one or more haplotypes in an individual and includes use of family pedigrees, molecular techniques and/or statistical inference.

Haplotype data - Information concerning one or more of the following for a specific gene: a listing of the haplotype pairs in each individual in a population; a listing of the different haplotypes in a population; frequency of each haplotype in that or other populations, and any known associations between one or more haplotypes and a trait.

Isoform – A particular form of a gene, mRNA, cDNA or the protein encoded thereby, distinguished from other forms by its particular sequence and/or structure.

Isogene – One of the isoforms of a gene found in a population. An isogene contains all of the polymorphisms present in the particular isoform of the gene.

Isolated – As applied to a biological molecule such as RNA, DNA, oligonucleotide, or protein, isolated means the molecule is substantially free of other biological molecules such as nucleic acids, proteins, lipids, carbohydrates, or other material such as cellular debris and growth media. Generally, the term "isolated" is not intended to refer to a complete absence of such material or to absence of water, buffers, or salts, unless they are present in amounts that substantially interfere with

the methods of the present invention.

Locus - A location on a chromosome or DNA molecule corresponding to a gene or a physical or phenotypic feature.

Naturally-occurring – A term used to designate that the object it is applied to, e.g., naturally-occurring polynucleotide or polypeptide, can be isolated from a source in nature and which has not been intentionally modified by man.

Nucleotide pair – The nucleotides found at a polymorphic site on the two copies of a chromosome from an individual.

Phased – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, phased means the combination of nucleotides present at those polymorphic sites on a single copy of the locus is known.

Polymorphic site (PS) – A position within a locus at which at least two alternative sequences are found in a population, the most frequent of which has a frequency of no more than 99%.

Polymorphic variant – A gene, mRNA, cDNA, polypeptide or peptide whose nucleotide or amino acid sequence varies from a reference sequence due to the presence of a polymorphism in the gene.

Polymorphism – The sequence variation observed in an individual at a polymorphic site. Polymorphisms include nucleotide substitutions, insertions, deletions and microsatellites and may, but need not, result in detectable differences in gene expression or protein function.

Polymorphism data – Information concerning one or more of the following for a specific gene: location of polymorphic sites; sequence variation at those sites; frequency of polymorphisms in one or more populations; the different genotypes and/or haplotypes determined for the gene; frequency of one or more of these genotypes and/or haplotypes in one or more populations; any known association(s) between a trait and a genotype or a haplotype for the gene.

Polymorphism Database – A collection of polymorphism data arranged in a systematic or methodical way and capable of being individually accessed by electronic or other means.

Polynucleotide – A nucleic acid molecule comprised of single-stranded RNA or DNA or comprised of complementary, double-stranded DNA.

Population Group – A group of individuals sharing a common ethnogeographic origin.

Reference Population – A group of subjects or individuals who are predicted to be representative of the genetic variation found in the general population. Typically, the reference population represents the genetic variation in the population at a certainty level of at least 85%, preferably at least 90%, more preferably at least 95% and even more preferably at least 99%.

Single Nucleotide Polymorphism (SNP) – Typically, the specific pair of nucleotides observed at a single polymorphic site. In rare cases, three or four nucleotides may be found.

Subject – A human individual whose genotypes or haplotypes or response to treatment or disease state are to be determined.

Treatment - A stimulus administered internally or externally to a subject.

Unphased – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, unphased means the combination of nucleotides present at those polymorphic sites on a single copy of the locus is not known.

The inventors herein have discovered 14 novel polymorphic sites in the IL13 gene. The polymorphic sites identified by the inventors are referred to as PS1-14 to designate the order in which they are located in the gene (see Table 2 below).

Thus, in one embodiment, the invention provides an isolated polynucleotide comprising a polymorphic variant of the IL13 gene or a fragment of the gene which contains at least one of the novel polymorphic sites described herein. The nucleotide sequence of a variant IL13 gene is identical to the reference genomic sequence for those portions of the gene examined, as described in the Examples below, except that it comprises a different nucleotide at one or more of the novel polymorphic sites PS1-PS14. Similarly, the nucleotide sequence of a variant fragment of the IL13 gene is identical to the corresponding portion of the reference sequence except for having a different nucleotide at one or more of the novel polymorphic sites described herein. Thus, the invention specifically does not include polynucleotides comprising a nucleotide sequence identical to the reference sequence (or other reported IL13 sequences) or to portions of the reference sequence (or other reported IL13 sequences), except for genotyping oligonucleotides as described below.

The location of a polymorphism in a variant gene or fragment is identified by aligning its sequence against SEQ ID NO:1. The polymorphism is selected from the group consisting of thymine at PS1, guanine at PS2, guanine at PS3, thymine at PS4, cytosine at PS5, adenine at PS6, thymine at PS7, adenine at PS8, thymine at PS9, thymine at PS10, adenine at PS11, thymine at PS12, thymine at PS13 and adenine at PS14. In a preferred embodiment, the polymorphic variant comprises a naturally-occurring isogene of the IL13 gene which is defined by any one of haplotypes 1-21 shown in Table 4 below.

Polymorphic variants of the invention may be prepared by isolating a clone containing the IL13 gene from a human genomic library. The clone may be sequenced to determine the identity of the nucleotides at the polymorphic sites described herein. Any particular variant claimed herein could be prepared from this clone by performing *in vitro* mutagenesis using procedures well-known in the art.

IL13 isogenes may be isolated using any method that allows separation of the two "copies" of the IL13 gene present in an individual, which, as readily understood by the skilled artisan, may be the same allele or different alleles. Separation methods include targeted *in vivo* cloning (TIVC) in yeast as described in WO 98/01573, U.S. Patent No. 5,866,404, and U.S. Patent No. 5,972,614. Another method, which is described in U.S. Patent No. 5,972,614, uses an allele specific oligonucleotide in combination with primer extension and exonuclease degradation to generate hemizygous DNA targets. Yet other methods are single molecule dilution (SMD) as described in Ruaño et al., Proc. Natl. Acad.

Sci. 87:6296-6300, 1990; and allele specific PCR (Ruaño et al., 17 Nucleic Acids. Res. 8392, 1989; Ruaño et al., 19 Nucleic Acids Res. 6877-6882, 1991; Michalatos-Beloin et al., 24 Nucleic Acids Res. 4841-4843, 1996).

The invention also provides IL13 genome anthologies, which are collections of IL13 isogenes found in a given population. The population may be any group of at least two individuals, including but not limited to a reference population, a population group, a family population, a clinical population, and a same sex population. An IL13 genome anthology may comprise individual IL13 isogenes stored in separate containers such as microtest tubes, separate wells of a microtitre plate and the like. Alternatively, two or more groups of the IL13 isogenes in the anthology may be stored in separate containers. Individual isogenes or groups of isogenes in a genome anthology may be stored in any convenient and stable form, including but not limited to in buffered solutions, as DNA precipitates, freeze-dried preparations and the like. A preferred IL13 genome anthology of the invention comprises a set of isogenes defined by the haplotypes shown in Table 4 below.

An isolated polynucleotide containing a polymorphic variant nucleotide sequence of the invention may be operably linked to one or more expression regulatory elements in a recombinant expression vector capable of being propagated and expressing the encoded IL13 protein in a prokaryotic or a eukaryotic host cell. Examples of expression regulatory elements which may be used include, but are not limited to, the lac system, operator and promoter regions of phage lambda, yeast promoters, and promoters derived from vaccinia virus, adenovirus, retroviruses, or SV40. Other regulatory elements include, but are not limited to, appropriate leader sequences, termination codons, polyadenylation signals, and other sequences required for the appropriate transcription and subsequent translation of the nucleic acid sequence in a given host cell. Of course, the correct combinations of expression regulatory elements will depend on the host system used. In addition, it is understood that the expression vector contains any additional elements necessary for its transfer to and subsequent replication in the host cell. Examples of such elements include, but are not limited to, origins of replication and selectable markers. Such expression vectors are commercially available or are readily constructed using methods known to those in the art (e.g., F. Ausubel et al., 1987, in "Current Protocols in Molecular Biology", John Wiley and Sons, New York, New York). Host cells which may be used to express the variant IL13 sequences of the invention include, but are not limited to, eukaryotic and mammalian cells, such as animal, plant, insect and yeast cells, and prokaryotic cells, such as E. coli, or algal cells as known in the art. The recombinant expression vector may be introduced into the host cell using any method known to those in the art including, but not limited to, microinjection, electroporation, particle bombardment, transduction, and transfection using DEAE-dextran, lipofection, or calcium phosphate (see e.g., Sambrook et al. (1989) in "Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Press, Plainview, New York). In a preferred aspect, eukaryotic expression vectors that function in eukaryotic cells, and preferably mammalian cells, are used. Non-limiting examples of such vectors include vaccinia virus vectors, adenovirus vectors,

herpes virus vectors, and baculovirus transfer vectors. Preferred eukaryotic cell lines include COS cells, CHO cells, HeLa cells, NIH/3T3 cells, and embryonic stem cells (Thomson, J. A. et al., 1998 Science 282:1145-1147). Particularly preferred host cells are mammalian cells.

As will be readily recognized by the skilled artisan, expression of polymorphic variants of the IL13 gene will produce IL13 mRNAs varying from each other at any polymorphic site retained in the spliced and processed mRNA molecules. These mRNAs can be used for the preparation of an IL13 cDNA comprising a nucleotide sequence which is a polymorphic variant of the IL13 reference coding sequence shown in Figure 2. Thus, the invention also provides IL13 mRNAs and corresponding cDNAs which comprise a nucleotide sequence that is identical to SEQ ID NO:2 (Fig. 2), or its corresponding RNA sequence, except for having one or both polymorphisms selected from the group consisting of adenine at a position corresponding to nucleotide 111 and adenine at a position corresponding to nucleotide 386. Fragments of these variant mRNAs and cDNAs are included in the scope of the invention, provided they contain the novel polymorphisms described herein. The invention specifically excludes polynucleotides identical to previously identified and characterized IL13 cDNAs and fragments thereof. Polynucleotides comprising a variant RNA or DNA sequence may be isolated from a biological sample using well-known molecular biological procedures or may be chemically synthesized.

Genomic and cDNA fragments of the invention comprise at least one novel polymorphic site identified herein and have a length of at least 10 nucleotides and may range up to the full length of the gene. Preferably, a fragment according to the present invention is between 100 and 3000 nucleotides in length, and more preferably between 200 and 2000 nucleotides in length, and most preferably between 500 and 1000 nucleotides in length.

In describing the polymorphic sites identified herein, reference is made to the sense strand of the gene for convenience. However, as recognized by the skilled artisan, nucleic acid molecules containing the IL13 gene may be complementary double stranded molecules and thus reference to a particular site on the sense strand refers as well to the corresponding site on the complementary antisense strand. Thus, reference may be made to the same polymorphic site on either strand and an oligonucleotide may be designed to hybridize specifically to either strand at a target region containing the polymorphic site. Thus, the invention also includes single-stranded polynucleotides which are complementary to the sense strand of the IL13 genomic variants described herein.

Polynucleotides comprising a polymorphic gene variant or fragment may be useful for therapeutic purposes. For example, where a patient could benefit from expression, or increased expression, of a particular IL13 protein isoform, an expression vector encoding the isoform may be administered to the patient. The patient may be one who lacks the IL13 isogene encoding that isoform or may already have at least one copy of that isogene.

In other situations, it may be desirable to decrease or block expression of a particular IL13 isogene. Expression of an IL13 isogene may be turned off by transforming a targeted organ, tissue or

cell population with an expression vector that expresses high levels of untranslatable mRNA for the isogene. Alternatively, oligonucleotides directed against the regulatory regions (e.g., promoter, introns, enhancers, 3′ untranslated region) of the isogene may block transcription. Oligonucleotides targeting the transcription initiation site, e.g., between positions −10 and +10 from the start site are preferred. Similarly, inhibition of transcription can be achieved using oligonucleotides that base-pair with region(s) of the isogene DNA to form triplex DNA (see e.g., Gee et al. in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, N.Y., 1994). Antisense oligonucleotides may also be designed to block translation of IL13 mRNA transcribed from a particular isogene. It is also contemplated that ribozymes may be designed that can catalyze the specific cleavage of IL13 mRNA transcribed from a particular isogene.

The oligonucleotides may be delivered to a target cell or tissue by expression from a vector introduced into the cell or tissue *in vivo* or *ex vivo*. Alternatively, the oligonucleotides may be formulated as a pharmaceutical composition for administration to the patient. Oligoribonucleotides and/or oligodeoxynucleotides intended for use as antisense oligonucleotides may be modified to increase stability and half-life. Possible modifications include, but are not limited to phosphorothioate or 2′ O-methyl linkages, and the inclusion of nontraditional bases such as inosine and queosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytosine, guanine, thymine, and uracil which are not as easily recognized by endogenous nucleases.

The invention also provides an isolated polypeptide comprising a polymorphic variant of the reference IL13 amino acid sequence shown in Figure 3. The location of a variant amino acid in an IL13 polypeptide or fragment of the invention is identified by aligning its sequence against SEQ ID NO:3 (Fig. 3). An IL13 protein variant of the invention comprises an amino acid sequence identical to SEQ ID NO:3 except for having glutamine at a position corresponding to amino acid position 129. The invention specifically excludes amino acid sequences identical to those previously identified for IL13, including SEQ ID NO:3, and previously described fragments thereof. IL13 protein variants included within the invention comprise all amino acid sequences based on SEQ ID NO:3 and having glutamine at a position corresponding to amino acid position 129.

The invention also includes IL13 peptide variants, which are any fragments of an IL13 protein variant that contains glutamine at a position corresponding to amino acid position 129. An IL13 peptide variant is at least 6 amino acids in length and is preferably any number between 6 and 30 amino acids long, more preferably between 10 and 25, and most preferably between 15 and 20 amino acids long. Such IL13 peptide variants may be useful as antigens to generate antibodies specific for one of the above IL13 isoforms. In addition, the IL13 peptide variants may be useful in drug screening assays.

An IL13 variant protein or peptide of the invention may be prepared by chemical synthesis or by expressing one of the variant IL13 genomic and cDNA sequences as described above. Alternatively, the IL13 protein variant may be isolated from a biological sample of an individual

having an IL13 isogene which encodes the variant protein. Where the sample contains two different IL13 isoforms (i.e., the individual has different IL13 isogenes), a particular IL13 isoform of the invention can be isolated by immunoaffinity chromatography using an antibody which specifically binds to that particular IL13 isoform but does not bind to the other IL13 isoform.

The expressed or isolated IL13 protein may be detected by methods known in the art, including Coomassie blue staining, silver staining, and Western blot analysis using antibodies specific for the isoform of the IL13 protein as discussed further below. IL13 variant proteins can be purified by standard protein purification procedures known in the art, including differential precipitation, molecular sieve chromatography, ion-exchange chromatography, isoelectric focusing, gel electrophoresis, affinity and immunoaffinity chromatography and the like. (Ausubel et. al., 1987, In Current Protocols in Molecular Biology John Wiley and Sons, New York, New York). In the case of immunoaffinity chromatography, antibodies specific for a particular polymorphic variant may be used.

A polymorphic variant IL13 gene of the invention may also be fused in frame with a heterologous sequence to encode a chimeric IL13 protein. The non-IL13 portion of the chimeric protein may be recognized by a commercially available antibody. In addition, the chimeric protein may also be engineered to contain a cleavage site located between the IL13 and non-IL13 portions so that the IL13 protein may be cleaved and purified away from the non-IL13 portion.

An additional embodiment of the invention relates to using a novel IL13 protein isoform in any of a variety of drug screening assays. Such screening assays may be performed to identify agents that bind specifically to all known IL13 protein isoforms or to only a subset of one or more of these isoforms. The agents may be from chemical compound libraries, peptide libraries and the like. The IL13 protein or peptide variant may be free in solution or affixed to a solid support. In one embodiment, high throughput screening of compounds for binding to an IL13 variant may be accomplished using the method described in PCT application WO84/03565, in which large numbers of test compounds are synthesized on a solid substrate, such as plastic pins or some other surface, contacted with the IL13 protein(s) of interest and then washed. Bound IL13 protein(s) are then detected using methods well-known in the art.

In another embodiment, a novel IL13 protein isoform may be used in assays to measure the binding affinities of one or more candidate drugs targeting the IL13 protein.

In another embodiment, the invention provides antibodies specific for and immunoreactive with one or more of the novel IL13 variant proteins described herein. The antibodies may be either monoclonal or polyclonal in origin. The IL13 protein or peptide variant used to generate the antibodies may be from natural or recombinant sources or produced by chemical synthesis using synthesis techniques known in the art. If the IL13 protein variant is of insufficient size to be antigenic, it may be conjugated, complexed, or otherwise covalently linked to a carrier molecule to enhance the antigenicity of the peptide. Examples of carrier molecules, include, but are not limited to, albumins

(e.g., human, bovine, fish, ovine), and keyhole limpet hemocyanin (Basic and Clinical Immunology, 1991, Eds. D.P. Stites, and A.I. Terr, Appleton and Lange, Norwalk Connecticut, San Mateo, California).

In one embodiment, an antibody specifically immunoreactive with one of the novel IL13 protein isoforms described herein is administered to an individual to neutralize activity of the IL13 isoform expressed by that individual. The antibody may be formulated as a pharmaceutical composition which includes a pharmaceutically acceptable carrier.

Antibodies specific for and immunoreactive with one of the novel IL13 protein isoform described herein may be used to immunoprecipitate the IL13 protein variant from solution as well as react with IL13 protein isoforms on Western or immunoblots of polyacrylamide gels on membrane supports or substrates. In another preferred embodiment, the antibodies will detect IL13 protein isoforms in paraffin or frozen tissue sections, or in cells which have been fixed or unfixed and prepared on slides, coverslips, or the like, for use in immunocytochemical, immunohistochemical, and immunofluorescence techniques.

In another embodiment, an antibody specifically immunoreactive with one of the novel IL13 protein variants described herein is used in immunoassays to detect this variant in biological samples. In this method, an antibody of the present invention is contacted with a biological sample and the formation of a complex between the IL13 protein variant and the antibody is detected. As described, suitable immunoassays include radioimmunoassay, Western blot assay, immunofluorescent assay, enzyme linked immunoassay (ELISA), chemiluminescent assay, immunohistochemical assay, immunocytochemical assay, and the like (see, e.g., Principles and Practice of Immunoassay, 1991, Eds. Christopher P. Price and David J. Neoman, Stockton Press, New York, New York; Current Protocols in Molecular Biology, 1987, Eds. Ausubel et al., John Wiley and Sons, New York, New York). Standard techniques known in the art for ELISA are described in Methods in Immunodiagnosis, 2nd Ed., Eds. Rose and Bigazzi, John Wiley and Sons, New York 1980; and Campbell et al., 1984, Methods in Immunology, W.A. Benjamin, Inc.). Such assays may be direct, indirect, competitive, or noncompetitive as described in the art (see, e.g., Principles and Practice of Immunoassay, 1991, Eds. Christopher P. Price and David J. Neoman, Stockton Pres, NY, NY; and Oellirich, M., 1984, J. Clin. Chem. Clin. Biochem., 22:895-904). Proteins may be isolated from test specimens and biological samples by conventional methods, as described in Current Protocols in Molecular Biology, supra.

Exemplary antibody molecules for use in the detection and therapy methods of the present invention are intact immunoglobulin molecules, substantially intact immunoglobulin molecules, or those portions of immunoglobulin molecules that contain the antigen binding site. Polyclonal or monoclonal antibodies may be produced by methods conventionally known in the art (e.g., Kohler and Milstein, 1975, Nature, 256:495-497; Campbell Monoclonal Antibody Technology, the Production and Characterization of Rodent and Human Hybridomas, 1985, In: Laboratory Techniques in

Biochemistry and Molecular Biology, Eds. Burdon et al., Volume 13, Elsevier Science Publishers, Amsterdam). The antibodies or antigen binding fragments thereof may also be produced by genetic engineering. The technology for expression of both heavy and light chain genes in E. coli is the subject of PCT patent applications, publication number WO 901443, WO 901443 and WO 9014424 and in Huse et al., 1989, Science, 246:1275-1281. The antibodies may also be humanized (e.g., Queen, C. et al. 1989 Proc. Natl. Acad. Sci. 86;10029).

Effect(s) of the polymorphisms identified herein on expression of IL13 may be investigated by preparing recombinant cells and/or organisms, preferably recombinant animals, containing a polymorphic variant of the IL13 gene. As used herein, "expression" includes but is not limited to one or more of the following: transcription of the gene into precursor mRNA; splicing and other processing of the precursor mRNA to produce mature mRNA; mRNA stability; translation of the mature mRNA into IL13 protein (including codon usage and tRNA availability); and glycosylation and/or other modifications of the translation product, if required for proper expression and function.

To prepare a recombinant cell of the invention, the desired IL13 isogene may be introduced into the cell in a vector such that the isogene remains extrachromosomal. In such a situation, the gene will be expressed by the cell from the extrachromosomal location. In a preferred embodiment, the IL13 isogene is introduced into a cell in such a way that it recombines with the endogenous IL13 gene present in the cell. Such recombination requires the occurrence of a double recombination event, thereby resulting in the desired IL13 gene polymorphism. Vectors for the introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, and any suitable vector or vector construct may be used in the invention. Methods such as electroporation, particle bombardment, calcium phosphate co-precipitation and viral transduction for introducing DNA into cells are known in the art; therefore, the choice of method may lie with the competence and preference of the skilled practitioner. Examples of cells into which the IL13 isogene may be introduced include, but are not limited to, continuous culture cells, such as COS, NIH/3T3, and primary or culture cells of the relevant tissue type, i.e., they express the IL13 isogene. Such recombinant cells can be used to compare the biological activities of the different protein variants.

Recombinant organisms, i.e., transgenic animals, expressing a variant IL13 gene are prepared using standard procedures known in the art. Preferably, a construct comprising the variant gene is introduced into a nonhuman animal or an ancestor of the animal at an embryonic stage, i.e., the one-cell stage, or generally not later than about the eight-cell stage. Transgenic animals carrying the constructs of the invention can be made by several methods known to those having skill in the art. One method involves transfecting into the embryo a retrovirus constructed to contain one or more insulator elements, a gene or genes of interest, and other components known to those skilled in the art to provide a complete shuttle vector harboring the insulated gene(s) as a transgene, see e.g., U.S. Patent No. 5,610,053. Another method involves directly injecting a transgene into the embryo. A third method involves the use of embryonic stem cells. Examples of animals into which the IL13

isogenes may be introduced include, but are not limited to, mice, rats, other rodents, and nonhuman primates (see "The Introduction of Foreign Genes into Mice" and the cited references therein, In: Recombinant DNA, Eds. J.D. Watson, M. Gilman, J. Witkowski, and M. Zoller; W.H. Freeman and Company, New York, pages 254-272). Transgenic animals stably expressing a human IL13 isogene and producing human IL13 protein can be used as biological models for studying diseases related to abnormal IL13 expression and/or activity, and for screening and assaying various candidate drugs, compounds, and treatment regimens to reduce the symptoms or effects of these diseases.

An additional embodiment of the invention relates to pharmaceutical compositions for treating disorders affected by expression or function of a novel IL13 isogene described herein. The pharmaceutical composition may comprise any of the following active ingredients: a polynucleotide comprising one of these novel IL13 isogenes; an antisense oligonucleotide directed against one of the novel IL13 isogenes, a polynucleotide encoding such an antisense oligonucleotide, or another compound which inhibits expression of a novel IL13 isogene described herein. Preferably, the composition contains the active ingredient in a therapeutically effective amount. By therapeutically effective amount is meant that one or more of the symptoms relating to disorders affected by expression or function of a novel IL13 isogene is reduced and/or eliminated. The composition also comprises a pharmaceutically acceptable carrier, examples of which include, but are not limited to, saline, buffered saline, dextrose, and water. Those skilled in the art may employ a formulation most suitable for the active ingredient, whether it is a polynucleotide, oligonucleotide, protein, peptide or small molecule antagonist. The pharmaceutical composition may be administered alone or in combination with at least one other agent, such as a stabilizing compound. Administration of the pharmaceutical composition may be by any number of routes including, but not limited to oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, intradermal, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, PA).

For any composition, determination of the therapeutically effective dose of active ingredient and/or the appropriate route of administration is well within the capability of those skilled in the art. For example, the dose can be estimated initially either in cell culture assays or in animal models. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans. The exact dosage will be determined by the practitioner, in light of factors relating to the patient requiring treatment, including but not limited to severity of the disease state, general health, age, weight and gender of the patient, diet, time and frequency of administration, other drugs being taken by the patient, and tolerance/response to the treatment.

Information on the identity of genotypes and haplotypes for the IL13 gene of any particular individual as well as the frequency of such genotypes and haplotypes in any particular population of

individuals is expected to be useful for a variety of basic research and clinical applications. Thus, the invention also provides compositions and methods for detecting the novel IL13 polymorphisms identified herein.

The compositions comprise at least one IL13 genotyping oligonucleotide. In one embodiment, an IL13 genotyping oligonucleotide is a probe or primer capable of hybridizing to a target region that is located close to, or that contains, one of the novel polymorphic sites described herein. As used herein, the term "oligonucleotide" refers to a polynucleotide molecule having less than about 100 nucleotides. A preferred oligonucleotide of the invention is 10 to 35 nucleotides long. More preferably, the oligonucleotide is between 15 and 30, and most preferably, between 20 and 25 nucleotides in length. The oligonucleotide may be comprised of any phosphorylation state of ribonucleotides, deoxyribonucleotides, and acyclic nucleotide derivatives, and other functionally equivalent derivatives. Alternatively, oligonucleotides may have a phosphate-free backbone, which may be comprised of linkages such as carboxymethyl, acetamidate, carbamate, polyamide (peptide nucleic acid (PNA)) and the like (Varma, R. in Molecular Biology and Biotechnology, A Comprehensive Desk Reference, Ed. R. Meyers, VCH Publishers, Inc. (1995), pages 617-620). Oligonucleotides of the invention may be prepared by chemical synthesis using any suitable methodology known in the art, or may be derived from a biological sample, for example, by restriction digestion. The oligonucleotides may be labeled, according to any technique known in the art, including use of radiolabels, fluorescent labels, enzymatic labels, proteins, haptens, antibodies, sequence tags and the like.

Genotyping oligonucleotides of the invention must be capable of specifically hybridizing to a target region of an IL13 polynucleotide, i.e., an IL13 isogene. As used herein, specific hybridization means the oligonucleotide forms an anti-parallel double-stranded structure with the target region under certain hybridizing conditions, while failing to form such a structure when incubated with a non-target region or a non-IL13 polynucleotide under the same hybridizing conditions. Preferably, the oligonucleotide specifically hybridizes to the target region under conventional high stringency conditions. The skilled artisan can readily design and test oligonucleotide probes and primers suitable for detecting polymorphisms in the IL13 gene using the polymorphism information provided herein in conjunction with the known sequence information for the IL13 gene and routine techniques.

A nucleic acid molecule such as an oligonucleotide or polynucleotide is said to be a "perfect" or "complete" complement of another nucleic acid molecule if every nucleotide of one of the molecules is complementary to the nucleotide at the corresponding position of the other molecule. A nucleic acid molecule is "substantially complementary" to another molecule if it hybridizes to that molecule with sufficient stability to remain in a duplex form under conventional low-stringency conditions. Conventional hybridization conditions are described, for example, by Sambrook J. et al., in Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989) and by Haymes, B.D. et al. in Nucleic Acid Hybridization, A Practical Approach,

IRL Press, Washington, D.C. (1985). While perfectly complementary oligonucleotides are preferred for detecting polymorphisms, departures from complete complementarity are contemplated where such departures do not prevent the molecule from specifically hybridizing to the target region. For example, an oligonucleotide primer may have a non-complementary fragment at its 5′ end, with the remainder of the primer being complementary to the target region. Alternatively, non-complementary nucleotides may be interspersed into the oligonucleotide probe or primer as long as the resulting probe or primer is still capable of specifically hybridizing to the target region.

Preferred genotyping oligonucleotides of the invention are allele-specific oligonucleotides. As used herein, the term allele-specific oligonucleotide (ASO) means an oligonucleotide that is able, under sufficiently stringent conditions, to hybridize specifically to one allele of a gene, or other locus, at a target region containing a polymorphic site while not hybridizing to the corresponding region in another allele(s). As understood by the skilled artisan, allele-specificity will depend upon a variety of readily optimized stringency conditions, including salt and formamide concentrations, as well as temperatures for both the hybridization and washing steps. Examples of hybridization and washing conditions typically used for ASO probes are found in Kogan et al., "Genetic Prediction of Hemophilia A" in PCR Protocols, A Guide to Methods and Applications, Academic Press, 1990 and Ruano et al., 87 Proc. Natl. Acad. Sci. USA 6296-6300, 1990. Typically, an allele-specific oligonucleotide will be perfectly complementary to one allele while containing a single mismatch for another allele.

Allele-specific oligonucleotide probes which usually provide good discrimination between different alleles are those in which a central position of the oligonucleotide probe aligns with the polymorphic site in the target region (e.g., approximately the 7th or 8th position in a 15 mer, the 8th or 9th position in a 16mer, the 10th or 11th position in a 20 mer). A preferred ASO probe for detecting IL13 gene polymorphisms comprises a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

```
AGGAAAACGAGGGAA   (SEQ ID NO:4) and its complement,
AGGAAAATGAGGGAA   (SEQ ID NO:5) and its complement,
ACAAAACAAAAAAAC   (SEQ ID NO:6) and its complement,
ACAAAACGAAAAAAC   (SEQ ID NO:7) and its complement,
CCTGGAAATCTGAAC   (SEQ ID NO:8) and its complement,
CCTGGAAGTCTGAAC   (SEQ ID NO:9) and its complement,
TGCCTGCCGCTGGAA   (SEQ ID NO:10) and its complement,
TGCCTGCTGCTGGAA   (SEQ ID NO:11) and its complement,
CCCAGCCTATGCATC   (SEQ ID NO:12) and its complement,
CCCAGCCCATGCATC   (SEQ ID NO:13) and its complement,
AGCTGGTCAACATCA   (SEQ ID NO:14) and its complement,
AGCTGGTAAACATCA   (SEQ ID NO:15) and its complement,
ACCAAGGCCCGGCCC   (SEQ ID NO:16) and its complement,
ACCAAGGTCCGGCCC   (SEQ ID NO:17) and its complement,
TGGCCTGGGCTGCCA   (SEQ ID NO:18) and its complement,
TGGCCTGAGCTGCCA   (SEQ ID NO:19) and its complement,
CTGCCAGCACTCTGC   (SEQ ID NO:20) and its complement,
CTGCCAGTACTCTGC   (SEQ ID NO:21) and its complement,
```

GAATATCCATGGTGT (SEQ ID NO:22) and its complement,
GAATATCTATGGTGT (SEQ ID NO:23) and its complement,
TTTGTGCGAGTCGTC (SEQ ID NO:24) and its complement,
TTTGTGCAAGTCGTC (SEQ ID NO:25) and its complement,
CCTCTGGCGTTCTAC (SEQ ID NO:26) and its complement,
CCTCTGGTGTTCTAC (SEQ ID NO:27) and its complement,
CTACTCACGTGCTGA (SEQ ID NO:28) and its complement,
CTACTCATGTGCTGA (SEQ ID NO:29) and its complement,
GAGGGACGGTTCAAC (SEQ ID NO:30) and its complement, and
GAGGGACAGTTCAAC (SEQ ID NO:31) and its complement.

An allele-specific oligonucleotide primer of the invention has a 3′ terminal nucleotide, or preferably a 3′ penultimate nucleotide, that is complementary to only one nucleotide of a particular SNP, thereby acting as a primer for polymerase-mediated extension only if the allele containing that nucleotide is present. Allele-specific oligonucleotide primers hybridizing to either the coding or noncoding strand are contemplated by the invention. A preferred ASO primer for detecting IL13 gene polymorphisms comprises a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

ACTTCTAGGAAAACG (SEQ ID NO:32); CTGCTCTTCCCTCGT (SEQ ID NO:33);
ACTTCTAGGAAAATG (SEQ ID NO:34); CTGCTCTTCCCTCAT (SEQ ID NO:35);
AACAAAACAAAACAA (SEQ ID NO:36); TTTGGTGTTTTTTTG (SEQ ID NO:37);
AACAAAACAAAACGA (SEQ ID NO:38); TTTGGTGTTTTTTCG (SEQ ID NO:39);
TGGAGACCTGGAAAT (SEQ ID NO:40); GTCAAAGTTCAGATT (SEQ ID NO:41);
TGGAGACCTGGAAGT (SEQ ID NO:42); GTCAAAGTTCAGACT (SEQ ID NO:43);
GCCCCTTGCCTGCCG (SEQ ID NO:44); GTTTAATTCCAGCGG (SEQ ID NO:45);
GCCCCTTGCCTGCTG (SEQ ID NO:46); GTTTAATTCCAGCAG (SEQ ID NO:47);
AAGCCACCCAGCCTA (SEQ ID NO:48); GGAGCGGATGCATAG (SEQ ID NO:49);
AAGCCACCCAGCCCA (SEQ ID NO:50); GGAGCGGATGCATGG (SEQ ID NO:51);
TTGAGGAGCTGGTCA (SEQ ID NO:52); TCTGGGTGATGTTGA (SEQ ID NO:53);
TTGAGGAGCTGGTAA (SEQ ID NO:54); TCTGGGTGATGTTTA (SEQ ID NO:55);
CCATGGACCAAGGCC (SEQ ID NO:56); ATGGCTGGGCCGGGC (SEQ ID NO:57);
CCATGGACCAAGGTC (SEQ ID NO:58); ATGGCTGGGCCGGAC (SEQ ID NO:59);
CAGACCTGGCCTGGG (SEQ ID NO:60); CTGCCCTGGCAGCCC (SEQ ID NO:61);
CAGACCTGGCCTGAG (SEQ ID NO:62); CTGCCCTGGCAGCTC (SEQ ID NO:63);
CAACCCCTGCCAGCA (SEQ ID NO:64); CAGTGAGCAGAGTGC (SEQ ID NO:65);
CAACCCCTGCCAGTA (SEQ ID NO:66); CAGTGAGCAGAGTAC (SEQ ID NO:67);
CTGGCTGAATATCCA (SEQ ID NO:68); GGACACACACCATGG (SEQ ID NO:69);
CTGGCTGAATATCTA (SEQ ID NO:70); GGACACACACCATAG (SEQ ID NO:71);
CCCTGGTTTGTGCGA (SEQ ID NO:72); GGCCGGGACGACTCG (SEQ ID NO:73);
CCCTGGTTTGTGCAA (SEQ ID NO:74); GGCCGGGACGACTTG (SEQ ID NO:75);
TCCCGGCCTCTGGCG (SEQ ID NO:76); ACGTGAGTAGAACGC (SEQ ID NO:77);
TCCCGGCCTCTGGTG (SEQ ID NO:78); ACGTGAGTAGAACAC (SEQ ID NO:79);
GGCGTTCTACTCACG (SEQ ID NO:80); AAGAGGTCAGCACGT (SEQ ID NO:81);
GGCGTTCTACTCATG (SEQ ID NO:82); AAGAGGTCAGCACAT (SEQ ID NO:83);
TTTCGCGAGGGACGG (SEQ ID NO:84); GTTTCAGTTGAACCG (SEQ ID NO:85);
TTTCGCGAGGGACAG (SEQ ID NO:86); and GTTTCAGTTGAACTG (SEQ ID NO:87).

Other genotyping oligonucleotides of the invention hybridize to a target region located one to several nucleotides downstream of one of the novel polymorphic sites identified herein. Such oligonucleotides are useful in polymerase-mediated primer extension methods for detecting one of the novel polymorphisms described herein and therefore such genotyping oligonucleotides are referred to

herein as "primer-extension oligonucleotides". In a preferred embodiment, the 3′-terminus of a primer-extension oligonucleotide is a deoxynucleotide complementary to the nucleotide located immediately adjacent to the polymorphic site. A particularly preferred oligonucleotide primer for detecting IL13 gene polymorphisms by primer extension terminates in a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

TCTAGGAAAA	(SEQ ID NO:88);	CTCTTCCCTC	(SEQ ID NO:89);
AAAACAAAAC	(SEQ ID NO:90);	GGTGTTTTTT	(SEQ ID NO:91);
AGACCTGGAA	(SEQ ID NO:92);	AAAGTTCAGA	(SEQ ID NO:93);
CCTTGCCTGC	(SEQ ID NO:94);	TAATTCCAGC	(SEQ ID NO:95);
CCACCCAGCC	(SEQ ID NO:96);	GCGGATGCAT	(SEQ ID NO:97);
AGGAGCTGGT	(SEQ ID NO:98);	GGGTGATGTT	(SEQ ID NO:99);
TGGACCAAGG	(SEQ ID NO:100);	GCTGGGCCGG	(SEQ ID NO:101);
ACCTGGCCTG	(SEQ ID NO:102);	CCCTGGCAGC	(SEQ ID NO:103);
CCCCTGCCAG	(SEQ ID NO:104);	TGAGCAGAGT	(SEQ ID NO:105);
GCTGAATATC	(SEQ ID NO:106);	CACACACCAT	(SEQ ID NO:107);
TGGTTTGTGC	(SEQ ID NO:108);	CGGGACGACT	(SEQ ID NO:109);
CGGCCTCTGG	(SEQ ID NO:110);	TGAGTAGAAC	(SEQ ID NO:111);
GTTCTACTCA	(SEQ ID NO:112);	AGGTCAGCAC	(SEQ ID NO:113);
CGCGAGGGAC	(SEQ ID NO:114);	and TCAGTTGAAC	(SEQ ID NO:115).

In some embodiments, a composition contains two or more differently labeled genotyping oligonucleotides for simultaneously probing the identity of nucleotides at two or more polymorphic sites. It is also contemplated that primer compositions may contain two or more sets of allele-specific primer pairs to allow simultaneous targeting and amplification of two or more regions containing a polymorphic site.

IL13 genotyping oligonucleotides of the invention may also be immobilized on or synthesized on a solid surface such as a microchip, bead, or glass slide (see, e.g., WO 98/20020 and WO 98/20019). Such immobilized genotyping oligonucleotides may be used in a variety of polymorphism detection assays, including but not limited to probe hybridization and polymerase extension assays. Immobilized IL13 genotyping oligonucleotides of the invention may comprise an ordered array of oligonucleotides designed to rapidly screen a DNA sample for polymorphisms in multiple genes at the same time.

In another embodiment, the invention provides a kit comprising at least two genotyping oligonucleotides packaged in separate containers. The kit may also contain other components such as hybridization buffer (where the oligonucleotides are to be used as a probe) packaged in a separate container. Alternatively, where the oligonucleotides are to be used to amplify a target region, the kit may contain, packaged in separate containers, a polymerase and a reaction buffer optimized for primer extension mediated by the polymerase, such as PCR.

The above described oligonucleotide compositions and kits are useful in methods for genotyping and/or haplotyping the IL13 gene in an individual. As used herein, the terms "IL13 genotype" and "IL13 haplotype" mean the genotype or haplotype contains the nucleotide pair or nucleotide, respectively, that is present at one or more of the novel polymorphic sites described herein

and may optionally also include the nucleotide pair or nucleotide present at one or more additional polymorphic sites in the IL13 gene. The additional polymorphic sites may be currently known polymorphic sites or sites that are subsequently discovered.

One embodiment of the genotyping method involves isolating from the individual a nucleic acid mixture comprising the two copies of the IL13 gene, or a fragment thereof, that are present in the individual, and determining the identity of the nucleotide pair at one or more of the polymorphic sites selected from PS1-PS14 in the two copies to assign an IL13 genotype to the individual. As will be readily understood by the skilled artisan, the two "copies" of a gene in an individual may be the same allele or may be different alleles. In a particularly preferred embodiment, the genotyping method comprises determining the identity of the nucleotide pair at each of PS1-14.

Typically, the nucleic acid mixture is isolated from a biological sample taken from the individual, such as a blood sample or tissue sample. Suitable tissue samples include whole blood, semen saliva, tears, urine, fecal material, sweat, buccal, skin and hair. The nucleic acid mixture may be comprised of genomic DNA, mRNA, or cDNA and, in the latter two cases, the biological sample must be obtained from an organ in which the IL13 gene is expressed. Furthermore it will be understood by the skilled artisan that mRNA or cDNA preparations would not be used to detect polymorphisms located in introns or in 5′ and 3′ nontranscribed regions. If an IL13 gene fragment is isolated, it must contain the polymorphic site(s) to be genotyped.

One embodiment of the haplotyping method comprises isolating from the individual a nucleic acid molecule containing only one of the two copies of the IL13 gene, or a fragment thereof, that is present in the individual and determining in that copy the identity of the nucleotide at one or more of the polymorphic sites PS1-PS14 in that copy to assign an IL13 haplotype to the individual. The nucleic acid may be isolated using any method capable of separating the two copies of the IL13 gene or fragment such as one of the methods described above for preparing IL13 isogenes, with targeted *in vivo* cloning being the preferred approach. As will be readily appreciated by those skilled in the art, any individual clone will only provide haplotype information on one of the two IL13 gene copies present in an individual. If haplotype information is desired for the individual's other copy, additional IL13 clones will need to be examined. Typically, at least five clones should be examined to have more than a 90% probability of haplotyping both copies of the IL13 gene in an individual. In a particularly preferred embodiment, the nucleotide at each of PS1-14 is identified.

In a preferred embodiment, an IL13 haplotype pair is determined for an individual by identifying the phased sequence of nucleotides at one or more of the polymorphic sites selected from PS1-PS14 in each copy of the IL13 gene that is present in the individual. In a particularly preferred embodiment, the haplotyping method comprises identifying the phased sequence of nucleotides at each of PS1-14 in each copy of the IL13 gene. When haplotyping both copies of the gene, the identifying step is preferably performed with each copy of the gene being placed in separate containers. However, it is also envisioned that if the two copies are labeled with different tags, or are

otherwise separately distinguishable or identifiable, it could be possible in some cases to perform the method in the same container. For example, if first and second copies of the gene are labeled with different first and second fluorescent dyes, respectively, and an allele-specific oligonucleotide labeled with yet a third different fluorescent dye is used to assay the polymorphic site(s), then detecting a combination of the first and third dyes would identify the polymorphism in the first gene copy while detecting a combination of the second and third dyes would identify the polymorphism in the second gene copy.

In both the genotyping and haplotyping methods, the identity of a nucleotide (or nucleotide pair) at a polymorphic site(s) may be determined by amplifying a target region(s) containing the polymorphic site(s) directly from one or both copies of the IL13 gene, or fragment thereof, and the sequence of the amplified region(s) determined by conventional methods. It will be readily appreciated by the skilled artisan that only one nucleotide will be detected at a polymorphic site in individuals who are homozygous at that site, while two different nucleotides will be detected if the individual is heterozygous for that site. The polymorphism may be identified directly, known as positive-type identification, or by inference, referred to as negative-type identification. For example, where a SNP is known to be guanine and cytosine in a reference population, a site may be positively determined to be either guanine or cytosine for an individual homozygous at that site, or both guanine and cytosine, if the individual is heterozygous at that site. Alternatively, the site may be negatively determined to be not guanine (and thus cytosine/cytosine) or not cytosine (and thus guanine/guanine).

In addition, the identity of the allele(s) present at any of the novel polymorphic sites described herein may be indirectly determined by genotyping a polymorphic site not disclosed herein that is in linkage disequilibrium with the polymorphic site that is of interest. Two sites are said to be in linkage disequilibrium if the presence of a particular variant at one site enhances the predictability of another variant at the second site (Stevens, JC 1999, *Mol. Diag.* **4**: 309-17). Polymorphic sites in linkage disequilibrium with the presently disclosed polymorphic sites may be located in regions of the gene or in other genomic regions not examined herein. Genotyping of a polymorphic site in linkage disequilibrium with the novel polymorphic sites described herein may be performed by, but is not limited to, any of the above-mentioned methods for detecting the identity of the allele at a polymorphic site.

The target region(s) may be amplified using any oligonucleotide-directed amplification method, including but not limited to polymerase chain reaction (PCR) (U.S. Patent No. 4,965,188), ligase chain reaction (LCR) (Barany et al., Proc. Natl. Acad. Sci. USA 88:189-193, 1991; WO90/01069), and oligonucleotide ligation assay (OLA) (Landegren et al., Science 241:1077-1080, 1988). Oligonucleotides useful as primers or probes in such methods should specifically hybridize to a region of the nucleic acid that contains or is adjacent to the polymorphic site. Typically, the oligonucleotides are between 10 and 35 nucleotides in length and preferably, between 15 and 30 nucleotides in length. Most preferably, the oligonucleotides are 20 to 25 nucleotides long. The exact

length of the oligonucleotide will depend on many factors that are routinely considered and practiced by the skilled artisan.

Other known nucleic acid amplification procedures may be used to amplify the target region including transcription-based amplification systems (U.S. Patent No. 5,130,238; EP 329,822; U.S. Patent No. 5,169,766, WO89/06700) and isothermal methods (Walker et al., *Proc. Natl. Acad. Sci. USA* **89**:392-396, 1992).

A polymorphism in the target region may also be assayed before or after amplification using one of several hybridization-based methods known in the art. Typically, allele-specific oligonucleotides are utilized in performing such methods. The allele-specific oligonucleotides may be used as differently labeled probe pairs, with one member of the pair showing a perfect match to one variant of a target sequence and the other member showing a perfect match to a different variant. In some embodiments, more than one polymorphic site may be detected at once using a set of allele-specific oligonucleotides or oligonucleotide pairs. Preferably, the members of the set have melting temperatures within 5°C, and more preferably within 2°C, of each other when hybridizing to each of the polymorphic sites being detected.

Hybridization of an allele-specific oligonucleotide to a target polynucleotide may be performed with both entities in solution, or such hybridization may be performed when either the oligonucleotide or the target polynucleotide is covalently or noncovalently affixed to a solid support. Attachment may be mediated, for example, by antibody-antigen interactions, poly-L-Lys, streptavidin or avidin-biotin, salt bridges, hydrophobic interactions, chemical linkages, UV cross-linking baking, etc. Allele-specific oligonucleotides may be synthesized directly on the solid support or attached to the solid support subsequent to synthesis. Solid-supports suitable for use in detection methods of the invention include substrates made of silicon, glass, plastic, paper and the like, which may be formed, for example, into wells (as in 96-well plates), slides, sheets, membranes, fibers, chips, dishes, and beads. The solid support may be treated, coated or derivatized to facilitate the immobilization of the allele-specific oligonucleotide or target nucleic acid.

The genotype or haplotype for the IL13 gene of an individual may also be determined by hybridization of a nucleic sample containing one or both copies of the gene to nucleic acid arrays and subarrays such as described in WO 95/11995. The arrays would contain a battery of allele-specific oligonucleotides representing each of the polymorphic sites to be included in the genotype or haplotype.

The identity of polymorphisms may also be determined using a mismatch detection technique, including but not limited to the RNase protection method using riboprobes (Winter et al., Proc. Natl. Acad. Sci. USA 82:7575, 1985; Meyers et al., Science 230:1242, 1985) and proteins which recognize nucleotide mismatches, such as the E. coli mutS protein (Modrich, P. Ann. Rev. Genet. 25:229-253, 1991). Alternatively, variant alleles can be identified by single strand conformation polymorphism (SSCP) analysis (Orita et al., Genomics 5:874-879, 1989; Humphries et al., in Molecular Diagnosis of

Genetic Diseases, R. Elles, ed., pp. 321-340, 1996) or denaturing gradient gel electrophoresis (DGGE) (Wartell et al., Nucl. Acids Res. 18:2699-2706, 1990; Sheffield et al., Proc. Natl. Acad. Sci. USA 86:232-236, 1989).

A polymerase-mediated primer extension method may also be used to identify the polymorphism(s). Several such methods have been described in the patent and scientific literature and include the "Genetic Bit Analysis" method (WO92/15712) and the ligase/polymerase mediated genetic bit analysis (U.S. Patent 5,679,524. Related methods are disclosed in WO91/02087, WO90/09455, WO95/17676, U.S. Patent Nos. 5,302,509, and 5,945,283. Extended primers containing a polymorphism may be detected by mass spectrometry as described in U.S. Patent No. 5,605,798. Another primer extension method is allele-specific PCR (Ruaño et al., Nucl. Acids Res. 17:8392, 1989; Ruaño et al., Nucl. Acids Res. 19, 6877-6882, 1991; WO 93/22456; Turki et al., J. Clin. Invest. 95:1635-1641, 1995). In addition, multiple polymorphic sites may be investigated by simultaneously amplifying multiple regions of the nucleic acid using sets of allele-specific primers as described in Wallace et al. (WO89/10414).

In another aspect of the invention, an individual's IL13 haplotype pair is predicted from its IL13 genotype using information on haplotype pairs known to exist in a reference population. In its broadest embodiment, the haplotyping prediction method comprises identifying an IL13 genotype for the individual at two or more polymorphic sites selected from PS1-PS14, enumerating all possible haplotype pairs which are consistent with the genotype, accessing data containing IL13 haplotype pairs identified in a reference population, and assigning a haplotype pair to the individual that is consistent with the data. In one embodiment, the reference haplotype pairs include the IL13 haplotype pairs shown in Table 3.

Generally, the reference population should be composed of randomly-selected individuals representing the major ethnogeographic groups of the world. A preferred reference population for use in the methods of the present invention comprises an approximately equal number of individuals from Caucasian, African American, Asian and Hispanic-Latino population groups with the minimum number of each group being chosen based on how rare a haplotype one wants to be guaranteed to see. For example, if one wants to have a q% chance of not missing a haplotype that exists in the population at a p% frequency of occurring in the reference population, the number of individuals (n) who must be sampled is given by $2n=\log(1-q)/\log(1-p)$ where p and q are expressed as fractions. A preferred reference population allows the detection of any haplotype whose frequency is at least 10% with about 99% certainty and comprises about 20 unrelated individuals from each of the four population groups named above. A particularly preferred reference population includes a 3-generation family representing one or more of the four population groups to serve as controls for checking quality of haplotyping procedures.

In a preferred embodiment, the haplotype frequency data for each ethnogeographic group is examined to determine whether it is consistent with Hardy-Weinberg equilibrium. Hardy-Weinberg

equilibrium (D.L. Hartl et al., Principles of Population Genomics, Sinauer Associates (Sunderland, MA), 3rd Ed., 1997) postulates that the frequency of finding the haplotype pair H_1 / H_2 is equal to $p_{H-W}(H_1 / H_2) = 2p(H_1)p(H_2)$ if $H_1 \neq H_2$ and $p_{H-W}(H_1 / H_2) = p(H_1)p(H_2)$ if $H_1 = H_2$. A statistically significant difference between the observed and expected haplotype frequencies could be due to one or more factors including significant inbreeding in the population group, strong selective pressure on the gene, sampling bias, and/or errors in the genotyping process. If large deviations from Hardy-Weinberg equilibrium are observed in an ethnogeographic group, the number of individuals in that group can be increased to see if the deviation is due to a sampling bias. If a larger sample size does not reduce the difference between observed and expected haplotype pair frequencies, then one may wish to consider haplotyping the individual using a direct haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), SMD, or allele-specific long-range PCR (Michalotos-Beloin et al., Nucleic Acids Res. 24:4841-4843, 1996).

In one embodiment of this method for predicting an IL13 haplotype pair, the assigning step involves performing the following analysis. First, each of the possible haplotype pairs is compared to the haplotype pairs in the reference population. Generally, only one of the haplotype pairs in the reference population matches a possible haplotype pair and that pair is assigned to the individual. Occasionally, only one haplotype represented in the reference haplotype pairs is consistent with a possible haplotype pair for an individual, and in such cases the individual is assigned a haplotype pair containing this known haplotype and a new haplotype derived by subtracting the known haplotype from the possible haplotype pair. In rare cases, either no haplotypes in the reference population are consistent with the possible haplotype pairs, or alternatively, multiple reference haplotype pairs are consistent with the possible haplotype pairs. In such cases, the individual is preferably haplotyped using a direct molecular haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), SMD, or allele-specific long-range PCR (Michalotos-Beloin et al., Nucleic Acids Res. 24:4841-4843, 1996).

The invention also provides a method for determining the frequency of an IL13 genotype or IL13 haplotype in a population. The method comprises determining the genotype or the haplotype pair for the IL13 gene that is present in each member of the population, wherein the genotype or haplotype comprises the nucleotide pair or nucleotide detected at one or more of the polymorphic sites PS1-PS14 in the IL13 gene; and calculating the frequency any particular genotype or haplotype is found in the population. The population may be a reference population, a family population, a same sex population, a population group, a trait population (e.g., a group of individuals exhibiting a trait of interest such as a medical condition or response to a therapeutic treatment).

In another aspect of the invention, frequency data for IL13 genotypes and/or haplotypes found in a reference population are used in a method for identifying an association between a trait and an IL13 genotype or an IL13 haplotype. The trait may be any detectable phenotype, including but not

limited to susceptibility to a disease or response to a treatment. The method involves obtaining data on the frequency of the genotype(s) or haplotype(s) of interest in a reference population as well as in a population exhibiting the trait. Frequency data for one or both of the reference and trait populations may be obtained by genotyping or haplotyping each individual in the populations using one of the methods described above. The haplotypes for the trait population may be determined directly or, alternatively, by the predictive genotype to haplotype approach described above. In another embodiment, the frequency data for the reference and/or trait populations is obtained by accessing previously determined frequency data, which may be in written or electronic form. For example, the frequency data may be present in a database that is accessible by a computer. Once the frequency data is obtained, the frequencies of the genotype(s) or haplotype(s) of interest in the reference and trait populations are compared. In a preferred embodiment, the frequencies of all genotypes and/or haplotypes observed in the populations are compared. If a particular genotype or haplotype for the IL13 gene is more frequent in the trait population than in the reference population at a statistically significant amount, then the trait is predicted to be associated with that IL13 genotype or haplotype. Preferably, the IL13 genotype or haplotype being compared in the trait and reference populations is selected from the full-genotypes and full-haplotypes shown in Tables 3 and 4, respectively, or from sub-genotypes and sub-haplotypes derived from these genotypes and haplotypes.

In a preferred embodiment of the method, the trait of interest is a clinical response exhibited by a patient to some therapeutic treatment, for example, response to a drug targeting IL13 or response to a therapeutic treatment for a medical condition. As used herein, "medical condition" includes but is not limited to any condition or disease manifested as one or more physical and/or psychological symptoms for which treatment is desirable, and includes previously and newly identified diseases and other disorders. As used herein the term "clinical response" means any or all of the following: a quantitative measure of the response, no response, and adverse response (i.e., side effects).

In order to deduce a correlation between clinical response to a treatment and an IL13 genotype or haplotype, it is necessary to obtain data on the clinical responses exhibited by a population of individuals who received the treatment, hereinafter the "clinical population". This clinical data may be obtained by analyzing the results of a clinical trial that has already been run and/or the clinical data may be obtained by designing and carrying out one or more new clinical trials. As used herein, the term "clinical trial" means any research study designed to collect clinical data on responses to a particular treatment, and includes but is not limited to phase I, phase II and phase III clinical trials. Standard methods are used to define the patient population and to enroll subjects.

It is preferred that the individuals included in the clinical population have been graded for the existence of the medical condition of interest. This is important in cases where the symptom(s) being presented by the patients can be caused by more than one underlying condition, and where treatment of the underlying conditions are not the same. An example of this would be where patients experience breathing difficulties that are due to either asthma or respiratory infections. If both sets were treated

with an asthma medication, there would be a spurious group of apparent non-responders that did not actually have asthma. These people would affect the ability to detect any correlation between haplotype and treatment outcome. This grading of potential patients could employ a standard physical exam or one or more lab tests. Alternatively, grading of patients could use haplotyping for situations where there is a strong correlation between haplotype pair and disease susceptibility or severity.

The therapeutic treatment of interest is administered to each individual in the trial population and each individual's response to the treatment is measured using one or more predetermined criteria. It is contemplated that in many cases, the trial population will exhibit a range of responses and that the investigator will choose the number of responder groups (e.g., low, medium, high) made up by the various responses. In addition, the IL13 gene for each individual in the trial population is genotyped and/or haplotyped, which may be done before or after administering the treatment.

After both the clinical and polymorphism data have been obtained, correlations between individual response and IL13 genotype or haplotype content are created. Correlations may be produced in several ways. In one method, individuals are grouped by their IL13 genotype or haplotype (or haplotype pair) (also referred to as a polymorphism group), and then the averages and standard deviations of clinical responses exhibited by the members of each polymorphism group are calculated.

These results are then analyzed to determine if any observed variation in clinical response between polymorphism groups is statistically significant. Statistical analysis methods which may be used are described in L.D. Fisher and G. vanBelle, "Biostatistics: A Methodology for the Health Sciences", Wiley-Interscience (New York) 1993. This analysis may also include a regression calculation of which polymorphic sites in the IL13 gene give the most significant contribution to the differences in phenotype. One regression model useful in the invention is described in the PCT Application entitled "Methods for Obtaining and Using Haplotype Data", filed June 26, 2000.

A second method for finding correlations between IL13 haplotype content and clinical responses uses predictive models based on error-minimizing optimization algorithms. One of many possible optimization algorithms is a genetic algorithm (R. Judson, "Genetic Algorithms and Their Uses in Chemistry" in Reviews in Computational Chemistry, Vol. 10, pp. 1-73, K. B. Lipkowitz and D. B. Boyd, eds. (VCH Publishers, New York, 1997). Simulated annealing (Press et al., "Numerical Recipes in C: The Art of Scientific Computing", Cambridge University Press (Cambridge) 1992, Ch. 10), neural networks (E. Rich and K. Knight, "Artificial Intelligence", 2nd Edition (McGraw-Hill, New York, 1991, Ch. 18), standard gradient descent methods (Press et al., supra Ch. 10), or other global or local optimization approaches (see discussion in Judson, supra) could also be used. Preferably, the correlation is found using a genetic algorithm approach as described in PCT Application entitled "Methods for Obtaining and Using Haplotype Data", filed June 26, 2000.

Correlations may also be analyzed using analysis of variation (ANOVA) techniques to determine how much of the variation in the clinical data is explained by different subsets of the

polymorphic sites in the IL13 gene. As described in PCT Application entitled "Methods for Obtaining and Using Haplotype Data", filed June 26, 2000, ANOVA is used to test hypotheses about whether a response variable is caused by or correlated with one or more traits or variables that can be measured (Fisher and vanBelle, supra, Ch. 10).

From the analyses described above, a mathematical model may be readily constructed by the skilled artisan that predicts clinical response as a function of IL13 genotype or haplotype content. Preferably, the model is validated in one or more follow-up clinical trials designed to test the model.

The identification of an association between a clinical response and a genotype or haplotype (or haplotype pair) for the IL13 gene may be the basis for designing a diagnostic method to determine those individuals who will or will not respond to the treatment, or alternatively, will respond at a lower level and thus may require more treatment, i.e., a greater dose of a drug. The diagnostic method may take one of several forms: for example, a direct DNA test (i.e., genotyping or haplotyping one or more of the polymorphic sites in the IL13 gene), a serological test, or a physical exam measurement. The only requirement is that there be a good correlation between the diagnostic test results and the underlying IL13 genotype or haplotype that is in turn correlated with the clinical response. In a preferred embodiment, this diagnostic method uses the predictive haplotyping method described above.

Any or all analytical and mathematical operations involved in practicing the methods of the present invention may be implemented by a computer. In addition, the computer may execute a program that generates views (or screens) displayed on a display device and with which the user can interact to view and analyze large amounts of information relating to the IL13 gene and its genomic variation, including chromosome location, gene structure, and gene family, gene expression data, polymorphism data, genetic sequence data, and clinical data population data (e.g., data on ethnogeographic origin, clinical responses, genotypes, and haplotypes for one or more populations). The IL13 polymorphism data described herein may be stored as part of a relational database (e.g., an instance of an Oracle database or a set of ASCII flat files). These polymorphism data may be stored on the computer's hard drive or may, for example, be stored on a CD ROM or on one or more other storage devices accessible by the computer. For example, the data may be stored on one or more databases in communication with the computer via a network.

Preferred embodiments of the invention are described in the following examples. Other embodiments within the scope of the claims herein will be apparent to one skilled in the art from consideration of the specification or practice of the invention as disclosed herein. It is intended that the specification, together with the examples, be considered exemplary only, with the scope and spirit of the invention being indicated by the claims which follow the examples.

EXAMPLES

The Examples herein are meant to exemplify the various aspects of carrying out the invention and are not intended to limit the scope of the invention in any way. The Examples do not include detailed descriptions for conventional methods employed, such as in the performance of genomic DNA isolation, PCR and sequencing procedures. Such methods are well-known to those skilled in the art and are described in numerous publications, for example, Sambrook, Fritsch, and Maniatis, "Molecular Cloning: A Laboratory Manual", 2nd Edition, Cold Spring Harbor Laboratory Press, USA, (1989).

Example 1A

This example illustrates examination of various regions of the IL13 gene for polymorphic sites using DNA from Index Repository 1A.

Amplification of Target Regions

The following target regions of the IL13 gene were amplified using the PCR primer pairs listed below, with the sequences presented in the 5′ to 3′ direction and nucleotide positions shown for each region corresponding to Figure 1

```
Fragment 1
Forward Primer
960-979            AGCCTGAGCAGGCAGAGAGG (SEQ ID NO:116)
Reverse Primer
Complement of 1515-1496    CTTGGCCCCACTGTGAGAGG (SEQ ID NO:117)
PCR product  556 nt

Fragment 2
Forward Primer
1273-1295          CCAGACACGTGCACACTACTTCC (SEQ ID NO:118)
Reverse Primer
Complement of 1771-1750    CAACCCTGAGTACCTGGACAGC (SEQ ID NO:119)
PCR product  499 nt

Fragment 3
Forward Primer
1510-1531          GCCAAGGAGGAATTAGGCAAGC (SEQ ID NO:120)
Reverse Primer
Complement of 2218-2197    AGCTCCCTGAGGGCTGTAGAGG (SEQ ID NO:121)
PCR product  709 nt

Fragment 4
Forward Primer
1880-1903          CCCTGGACCCTTCTCAATAAGTCC (SEQ ID NO:122)
Reverse Primer
Complement of 2416-2396    CAGCTCCTCTCTCCCTCATGG (SEQ ID NO:123)
PCR product  537 nt
```

```
Fragment 5
Forward Primer
1923-1946        CCTTTATGCGACACTGGATTTTCC (SEQ ID NO:124)
Reverse Primer
Complement of 2495-2474    AGGCCAGGAGAGCAGGTAGTCC (SEQ ID NO:125)
PCR product  573 nt

Fragment 6
Forward Primer
3146-3168        CAAAGCAGAAAACGAAGCTCAGG (SEQ ID NO:126)
Reverse Primer
Complement of 3568-3546    TTTGTGAGTCTGCAGGAAGATGG (SEQ ID NO:127)
PCR product  423 nt

Fragment 7
Forward Primer
3451-3472        GTTCCAAGCAAGCTTCAAGTGC (SEQ ID NO:128)
Reverse Primer
Complement of 3914-3893    TGAGGCCCAGGTAGGAGTTAGG (SEQ ID NO:129)
PCR product  464 nt

Fragment 8
Forward Primer
3900-3920        CCTACCTGGGCCTCAATTTCC (SEQ ID NO:130)
Reverse Primer
Complement of 4527-4507    GTGTGAAGTGGGTCCCTGTCC (SEQ ID NO:131)
PCR product  628 nt
```

These primer pairs were used in PCR reactions containing genomic DNA isolated from immortalized cell lines for each member of the Index Repository. The PCR reactions were carried out under the following conditions:

Reaction volume	= 20 µl
10 x Advantage 2 Polymerase reaction buffer (Clontech)	= 2 µl
100 ng of human genomic DNA	= 1 µl
10 mM dNTP	= 0.4 µl
Advantage 2 Polymerase enzyme mix (Clontech)	= 0.2 µl
Forward Primer (10 µM)	= 0.4 µl
Reverse Primer (10 µM)	= 0.4 µl
Water	=15.6µl

Amplification profile:
94°C - 2 min. 1 cycle

94°C - 30 sec.
70°C - 45 sec.
72°C - 1 min.
} 10 cycles

94°C - 30 sec.
64°C - 45 sec.
72°C - 1 min.
} 35 cycles

Sequencing of PCR Products

The PCR products were purified by Solid Phase Reversible Immobilization using the protocol developed by the Whitehead Genome Center. A detailed protocol can be found at http://www.genome.wi.mit.edu/sequencing/protocols/pure/SPRI_pcr.html.

Briefly, five µl of carboxyl coated magnetic beads (10 mg/ml) and 60 µl of HYB BUFFER (2.5M NaCl/20% PEG 8000) were added to each PCR reaction mixture (20 µl). The reaction mixture was mixed well and incubated at room temperature (RT) for 10 min. The microtitre plate was placed on a magnet for 2 min and the beads washed twice with 150 µl of 70% EtOH. The beads were air dried for 2 min and the DNA was eluted in 25 µl of distilled water and incubated at RT for 5 min. The beads were magnetically separated and the supernatant removed for testing and sequencing.

The purified PCR products were sequenced in both directions using the primer sets described previously or those listed, in the 5′ to 3′ direction, below.

```
Fragment 1
Forward Primer
961-980            GCCTGAGCAGGCAGAGAGGG (SEQ ID NO:132)
Reverse Primer
Complement of 1514-1495    TTGGCCCCACTGTGAGAGGG (SEQ ID NO:133)

Fragment 2
Forward Primer
1275-1296          AGACACGTGCACACTACTTCCA (SEQ ID NO:134)
Reverse Primer
Complement of 1769-1749    ACCCTGAGTACCTGGACAGCT (SEQ ID NO:135)

Fragment 3
Forward Primer
1635-1655          CCCCAAATTCCCATAGCTGGT (SEQ ID NO:136)
Reverse Primer
Complement of 2152-2133    GCAATGACCGTGGTCAACAA (SEQ ID NO:137)

Fragment 4
Forward Primer
1930-1949          GCGACACTGGATTTTCCACA (SEQ ID NO:138)
Reverse Primer
Complement of 2387-2368    GGCCTTGGTCCATGGAGATT (SEQ ID NO:139)

Fragment 5
Forward Primer
1989-2010          GTGCCTGAAAAAGCAGAGACCA (SEQ ID NO:140)
Reverse Primer
Complement of 2447-2428    AGGTCCATCCACAGTGCTGA (SEQ ID NO:141)

Fragment 6
Forward Primer
3190-3208          CTGCCTCTGTGCCACACCA (SEQ ID NO:142)
Reverse Primer
Complement of 3493-3475    ATTATGGCGGGAGGGAGGA (SEQ ID NO:143)
```

```
Fragment 7
Forward Primer
3477-3496        CTCCCTCCCGCCATAATCTG (SEQ ID NO:144)
Reverse Primer
Complement of 3871-3852     TAAATCCCGAAGGCCACGTC (SEQ ID NO:145)

Fragment 8
Forward Primer
3949-3968        ACCCACCTCATGGGGACTTC (SEQ ID NO:146)
Reverse Primer
Complement of 4349-4330     GGTCGGCTAGGCTGAAGACG (SEQ ID NO:147)
```

Example 1B

This example illustrates examination of the IL13 gene for polymorphic sites in several regions spanning about 2100 base pairs upstream of the ATG start site to about 115 base pairs downstream of the termination codon.

Amplification of Target Regions

The locations of the PCR primer pairs, which were designed based on Figure 1 are set forth below:

```
Fragment 1 (Promoter region)
Forward primer
21-43                       GAAACAAGCAACAGGACCCTCTG (SEQ ID NO:148)
Reverse primer
complement of 622-598       GGAACTACCTTGGTCTGTGCAACTC (SEQ ID NO:149)
PCR Product  602 nt

Fragment 2 (Promoter region)
Forward primer:
560-581                GGGATAAGGGGCGTTGACTCAC (SEQ ID NO:150)
Reverse primer:
complement of 1248-1223   CCATCCTTCTGGGAAACTCTTAACTG (SEQ ID NO:151)
PCR Product  689 nt

Fragment 3  (Promoter region)
Forward primer:
1101-1076              GTTTCTGGAGGACTTCTAGGAAAACG (SEQ ID NO:152)
Reverse primer:
complement of 1667-1644     TGACCACAGTCTACCAGCTATGGG (SEQ ID NO:153)
PCR Product 592 nt

Fragment 4 (Promoter region)
Forward primer:
1361-1384                   GCAACATAGTGAGACCCCATCTCC (SEQ ID NO:154)
Reverse primer:
complement of 2058-2035     GCGTCTTGTGGCAGCTTTTATAGG (SEQ ID NO:155)
PCR Product  698 nt
```

```
Fragment 5 (Promoter region)
Forward primer:
1635-1657                     CCCCAAATTCCCATAGCTGGTAG (SEQ ID NO:156)
Reverse primer:
complement of 2267-2246       GCCGACACTCTCCTTCTGGTTC (SEQ ID NO:157)
PCR Product  633 nt

Fragment 6 (Exon 1)
Forward primer:
1985-2009                     GATAGTGCCTGAAAAAGCAGAGACC (SEQ ID NO:158)
Reverse primer:
complement of 2454-2432       CCTCCATAGGTCCATCCACAGTG (SEQ ID NO:159)
PCR Product 470 nt

Fragment 7 (Exon 2 and 3)
Forward primer:  3126-3146        TCACCCCTATGCCTGCTGTTC (SEQ ID NO:160
)
Reverse primer:
complement of 3906-3883       AGGTAGGAGTTAGGAGCCCTTGAG (SEQ ID NO:161)
PCR Product  781 nt

Fragment 8 (Exon 4)
Forward primer:
3982-4002                     TGAGACAGTCCCTGGAAAGCC (SEQ ID NO:162)
Reverse primer:
complement of 4304-4281       AAGGAATTTTACCCCTCCCTAACC (SEQ ID NO:163)
PCR Product 324 nt
```

These primer pairs were used in PCR reactions containing genomic DNA isolated from immortalized cell lines for each member of the Index Repository. The PCR reactions were carried out under the following conditions:

Reaction volume	= 20 µl
10 x Advantage 2 Polymerase reaction buffer (Clontech)	= 2 µl
100 ng of human genomic DNA	= 1 µl
10 mM dNTP	= 0.4 µl
Advantage 2 Polymerase enzyme mix (Clontech)	= 0.2 µl
Forward Primer (10 µM)	= 0.4 µl
Reverse Primer (10 µM)	= 0.4 µl
Water	=15.6µl

Amplification profile:
94°C - 2 min. 1 cycle

94°C - 30 sec.
70°C - 45 sec.
72°C - 1 min.
} 10 cycles

94°C - 30 sec.
64°C - 45 sec.
72°C - 1 min.
} 35 cycles

Sequencing of PCR Products

The PCR products were purified by Solid Phase Reversible Immobilization using the protocol developed by the Whitehead Genome Center. A detailed protocol can be found at http://www.genome.wi.mit.edu/sequencing/protocols/pure/SPRI_pcr.html.

Briefly, carboxyl coated magnetic beads (10 mg/ml) were washed three times with wash buffer (0.5 M EDTA, pH 8.0). Ten µl of washed beads and 50 µl of HYB BUFFER (2.5M NaCl/20% PEG 8000) were added to each PCR reaction mixture (50 µl). The reaction mixture was mixed well and incubated at room temperature (RT) for 10 min. The microtitre plate was placed on a magnet for 2 min and the beads washed twice with 150 µl of 70% EtOH. The beads were air dried for 2 min and resuspend in 20 µl of elution buffer (10 mM trisacetate, pH 7.8) and incubated at RT for 5 min. The beads were magnetically separated and the supernatant removed for testing and sequencing.

The purified PCR products were sequenced in both directions using the above PCR primer sets except where indicated below, with nucleotide positions of the primers referring to Fig. 1.

```
Fragment 1
Forward primer:
26-49                        AAGCAACAGGACCCTCTGATGTAG (SEQ ID NO:164)
Reverse primer:
complement of 620-597        AACTACCTTGGTCTGTGCAACTCC (SEQ ID NO:165)

Fragment 3
Forward primer:
1089-1111                    TTCTAGGAAAACGAGGGAAGAGC (SEQ ID NO:166)
Reverse primer:
complement of 1667-1644      TGACCACAGTCTACCAGCTATGGG (SEQ ID NO:167)

Fragment 4
Forward primer:
1443-1468               GAATCTTTCTGGATCTCTCAGTGGAG (SEQ ID NO:168)
Reverse primer:
complement of 2032-2012      CAAGTGGTGACGCCTGACTCC (SEQ ID NO:169)

Fragment 5
Forward primer:
1683-1707                    AAGACTATCTGCTCAGCACTTCTGG (SEQ ID NO:170)
Reverse primer:
complement of 2241-2219      TGATGTTGACCAGCTCCTCAATG (SEQ ID NO:171)

Fragment 6 (exon 1)
Forward primer:
2066-2085               ACAAGCCACCCAGCCTATGC (SEQ ID NO:172)
Reverse primer:
complement of 2442-2423 CATCCACAGTGCTGAGCCCCCCAGCCCAGC(SEQID NO:173)
```

```
Fragment 7 (exon 2)
Forward primer:
3202-3221                      CACACCAGGGATGCTTGTGG (SEQ ID NO:174)
Reverse primer:
complement of 3556-3536    CTGGAAGATGGGGCTGAGATG (SEQ ID NO:175)

Fragment 8 (exon 3)
Forward primer:
3336-3358       TCCCTGATCAACGTGTCAGGCTGCAGTGCCATC (SEQ ID NO:176)
Reverse primer:
complement of 3805-3784  TATTCAGCCAGCTTCCCTTCAG (SEQ ID NO:177)

Fragment 9 (exon 4)
Forward primer:
4003-4022            CCTGGTTTGTGCGAGTCGTC (SEQ ID NO:178)
Reverse primer:
complement of 4293-4274  CCCCTCCCTAACCCTCCTTC (SEQ ID NO:179)
```

Analysis of Sequences for Polymorphic Sites

Sequences were analyzed for the presence of polymorphisms using the Polyphred program (Nickerson et al., Nucleic Acids Res. 14:2745-2751, 1997). The presence of a polymorphism was confirmed on both strands. The polymorphisms and their locations in the IL13 gene are listed in Table 2 below.

Table 2. Polymorphic Sites Identified in the IL13 Gene

Polymorphic Site Number	Nucleotide Position	Reference Allele	Variant Allele	Example
PS1 [R]	1100	C	T	1A, 1B
PS2 [R]	1408	A	G	1A, 1B
PS3	1477	A	G	1A
PS4	1842	C	T	1A
PS5	2080	T	C	1A
PS6	2233	C	A	1A
PS7	2385	C	T	1A
PS8	3248	G	A	1A
PS9 [R]	3280	C	T	1A, 1B
PS10	3807	C	T	1A
PS11	4014	G	A	1A
PS12	4033	C	T	1A
PS13 [R]	4044	C	T	1A, 1B
PS14 [R]	4165	G	A	1A, 1B

R: These polymorphic sites have previously been mapped to nucleotide positions 1100, 1408, 3281, 4045 and 4166, respectively, in the GenBank Accession No. U10307.1 in Application No. 60,156,489 filed on September 28, 1999.

Example 2

This example illustrates analysis of the IL13 polymorphisms identified in the Index Repositories IA and IB for human genotypes and haplotypes.

The different genotypes containing these polymorphisms that were observed in the reference population are shown in Table 3 below, with the haplotype pair indicating the combination of haplotypes determined for the individual using the haplotype derivation protocol described below. In Table 3, homozygous positions are indicated by one nucleotide and heterozygous positions are indicated by two nucleotides. Missing nucleotides in any given genotype in Table 3 can typically be inferred based on linkage disequilibrium and/or Mendelian inheritance.

Table 3(Part1). Genotypes and Haplotype Pairs Observed for IL13 Gene

Genotype	Polymorphic Sites										
Number	PS1	PS2	PS3	PS4	PS5	PS6	PS7	PS8	PS9	PS10	HAP Pair
1	C	A	A	C	T	C	C	G	C	C	3 3
2	T	A	A	C	T	C	C	G	C	C	14 14
3	C	A	A	C	T	C	C	G	C	C	4 4
4	C	A	A	C	T	C	C	G	C	C	5 5
5	T	A	G	C	T	C	C	G	C	C	19 19
6	T	A	A	C	T	C	C	G	C	C	15 15
7	C	A	A	C	T	C	C	G	C	C	3 6
8	C	A	A	C	T	C	C	G	C	C	4 6
9	C/T	A	A	C/T	T	C	C	G	C	C	3 17
10	T	A	A/G	C	T	C	C	G	C	C	14 18
11	C	A	A	C	T	C	C/T	G	C	C	3 9
12	C	A/G	A	C	T	C	C	G	C	C	3 10
13	C/T	A	A/G	C	T	C	C	G	C	C	4 19
14	C/T	A	A/G	C	T	C	C	G	C/T	C	3 21
15	C	A	A	C	T	C	C	G	C	C	3 5
16	T	A	A	C	T	C/A	C	G	C	C/T	14 12
17	C/T	A	A	C	T	C/A	C	G	C	C	3 11
18	C	A	A	C	T	C	C	G	C	C/T	3 7
19	C	A	A	C	T	C	C	G	C	C/T	5 7
20	C/T	A	A/G	C	T	C	C	G	C	C	3 20
21	C	A	A	C	T	C	C	G	C	C	4 5
22	T/C	A	G/A	C	T	C	C	G	C	C/T	19 7
23	C/T	A	A	C	T	C	C	G	C	C	3 16
24	C/T	A	A	C	T	C	C	G	C	C	3 15
25	C	A	A	C	T	C	C	G/A	C	C	3 2
26	C/T	A	A/G	C	T	C	C	G	C	C	3 19
27	C	A	A	C	T/C	C	C	G	C	C	4 1
28	C	A	A	C	T	C	C	G	C	C	3 4
29	C/T	A	A/G	C	T	C	C	G	C	C	3 18
30	T/C	A	A	C	T	C	C	G	C	C	14 4
31	C	A	A	C	T	C	C	G	C/T	C	5 8
32	T/C	A	G/A	C	T	C	C	G	C	C	19 5
33	T/C	A	G/A	C	T	C	C	G	C/T	C	19 8
34	C/T	A	A	C	T	C	C	G	C	C	3 13
35	T	A	A/G	C	T	C	C	G	C	C	14 19
36	C/T	A	A	C	T	C	C	G	C	C	3 14

Table 3(Part2). Genotypes and Haplotype Pairs Observed for IL13 Gene

Genotype Number	Polymorphic Sites PS11	PS12	PS13	PS14	HAP	Pair
1	G	C	C	G	3	3
2	G	C	T	A	14	14
3	G	C	T	A	4	4
4	G	C	T	G	5	5
5	G	C	T	G	19	19
6	G	C	T	G	15	15
7	G	C/T	C	G	3	6
8	G	C/T	T/C	A/G	4	6
9	G	C	C/T	G	3	17
10	G	C	T/C	A/G	14	18
11	G	C	C/T	G/A	3	9
12	G	C	C/T	G	3	10
13	G	C	T	A/G	4	19
14	G/A	C	C/T	G	3	21
15	G	C	C/T	G	3	5
16	G	C	T/C	A/G	14	12
17	G	C	C/T	G/A	3	11
18	G	C	C/T	G/A	3	7
19	G	C	T	G/A	5	7
20	G	C/T	C/T	G	3	20
21	G	C	T	A/G	4	5
22	G	C	T	G/A	19	7
23	G	C/T	C	G	3	16
24	G	C	C/T	G	3	15
25	G	C	C/T	G/A	3	2
26	G	C	C/T	G	3	19
27	G	C	T	A	4	1
28	G	C	C/T	G/A	3	4
29	G	C	C	G	3	18
30	G	C	T	A	14	4
31	G	C	T/C	G	5	8
32	G	C	T	G	19	5
33	G	C	T/C	G	19	8
34	G	C	C	G	3	13
35	G	C	T	A/G	14	19
36	G	C	C/T	G/A	3	14

The haplotype pairs shown in Table 3 were estimated from the unphased genotypes using an extension of Clark's algorithm (Clark, A.G. (1990) *Mol Bio Evol* 7, 111-122), as described in U.S. Provisional Patent Application filed April 19, 2000 and entitled "A Method and System for Determining Haplotypes from a Collection of Polymorphisms". In this method, haplotypes are assigned directly from individuals who are homozygous at all sites or heterozygous at no more than one of the variable sites. This list of haplotypes is then used to deconvolute the unphased genotypes in the remaining (multiply heterozygous) individuals.

By following this protocol, it was determined that the Index Repositories examined herein and, by extension, the general population contains the 21 human IL13 haplotypes shown in Table 4 below.

Table 4. Haplotypes Identified in the IL13 Gene

Haplotype Number	Polymorphic Sites PS 1	PS 2	PS 3	PS 4	PS 5	PS 6	PS 7	PS 8	PS 9	PS 10	PS 11	PS 12	PS 13	PS 14
1	C	A	A	C	C	C	C	G	C	C	G	C	T	A
2	C	A	A	C	T	C	C	A	C	C	G	C	T	A
3	C	A	A	C	T	C	C	G	C	C	G	C	C	G
4	C	A	A	C	T	C	C	G	C	C	G	C	T	A
5	C	A	A	C	T	C	C	G	C	C	G	C	T	G
6	C	A	A	C	T	C	C	G	C	C	G	T	C	G
7	C	A	A	C	T	C	C	G	C	T	G	C	T	A
8	C	A	A	C	T	C	C	G	T	C	G	C	C	G
9	C	A	A	C	T	C	T	G	C	C	G	C	T	A
10	C	G	A	C	T	C	C	G	C	C	G	C	T	G
11	T	A	A	C	T	A	C	G	C	C	G	C	T	A
12	T	A	A	C	T	A	C	G	C	T	G	C	C	G
13	T	A	A	C	T	C	C	G	C	C	G	C	C	G
14	T	A	A	C	T	C	C	G	C	C	G	C	T	A
15	T	A	A	C	T	C	C	G	C	C	G	C	T	G
16	T	A	A	C	T	C	C	G	C	C	G	T	C	G
17	T	A	A	T	T	C	C	G	C	C	G	C	T	G
18	T	A	G	C	T	C	C	G	C	C	G	C	C	G
19	T	A	G	C	T	C	C	G	C	C	G	C	T	G
20	T	A	G	C	T	C	C	G	C	C	G	T	T	G
21	T	A	G	C	T	C	C	G	T	C	A	C	T	G

In view of the above, it will be seen that the several advantages of the invention are achieved and other advantageous results attained.

As various changes could be made in the above methods and compositions without departing from the scope of the invention, it is intended that all matter contained in the above description and shown in the accompanying drawings shall be interpreted as illustrative and not in a limiting sense.

All references cited in this specification, including patents and patent applications, are hereby incorporated in their entirety by reference. The discussion of references herein is intended merely to summarize the assertions made by their authors and no admission is made that any reference constitutes prior art. Applicants reserve the right to challenge the accuracy and pertinency of the cited references.

What is Claimed is:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:
   (a) a first nucleotide sequence which is a polymorphic variant of a reference sequence for interleukin 13 (IL13) gene or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:1, and the polymorphic variant comprises at least one polymorphism selected from the group consisting of thymine at PS1, guanine at PS2, guanine at PS3, thymine at PS4, cytosine at PS5, adenine at PS6, thymine at PS7, adenine at PS8, thymine at PS9, thymine at PS10, adenine at PS11, thymine at PS12, thymine at PS13 and adenine at PS14; and
   (b) a second nucleotide sequence which is complementary to the first nucleotide sequence.
2. The isolated polynucleotide of claim 1 which comprises an IL13 isogene.
3. The isolated polynucleotide of claim 1 which is a DNA molecule and comprises both the first and second nucleotide sequences and further comprises expression regulatory elements operably linked to the first nucleotide sequence.
4. A recombinant nonhuman organism transformed or transfected with the isolated polynucleotide of claim 1, wherein the organism expresses an IL13 protein encoded by the first nucleotide sequence.
5. The recombinant organism of claim 4 which is a nonhuman transgenic animal.
6. The isolated polynucleotide of claim 1, wherein the first nucleotide sequence is a polymorphic variant of a fragment of the IL13 gene, the fragment comprising one or more polymorphisms selected from the group consisting of thymine at PS1, guanine at PS2, guanine at PS3, thymine at PS4, cytosine at PS5, adenine at PS6, thymine at PS7, adenine at PS8, thymine at PS9, thymine at PS10, adenine at PS11, thymine at PS12, thymine at PS13 and adenine at PS14.
7. An isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the IL13 cDNA or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:2 and the polymorphic variant comprises at least one polymorphism selected from the group consisting of adenine at a position corresponding to nucleotide 111 and adenine at a position corresponding to nucleotide 386.
8. A recombinant nonhuman organism transformed or transfected with the isolated polynucleotide of claim 7, wherein the organism expresses a interleukin 13 (IL13) protein encoded by the polymorphic variant sequence.
9. The recombinant organism of claim 8 which is a nonhuman transgenic animal.
10. An isolated polypeptide comprising an amino acid sequence which is a polymorphic variant of a reference sequence for the IL13 protein or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:3 and the polymorphic variant comprises glutamine at a position corresponding to amino acid position 129.
11. An isolated antibody specific for and immunoreactive with the isolated polypeptide of claim 10.
12. A method for screening for drugs targeting the isolated polypeptide of claim 10 which comprises

contacting the IL13 polymorphic variant with a candidate agent and assaying for binding activity.

13. A composition comprising at least one genotyping oligonucleotide for detecting a polymorphism in the interleukin 13 (IL13) gene at a polymorphic site selected from PS1-PS14.
14. The composition of claim 13, wherein the genotyping oligonucleotide is an allele-specific oligonucleotide that specifically hybridizes to an allele of the IL13 gene at a region containing the polymorphic site.
15. The composition of claim 14, wherein the allele-specific oligonucleotide comprises a nucleotide sequence selected from the group consisting of of SEQ ID NOS:4-31, the complements of SEQ ID NOS: 4-31, and SEQ ID NOS:32-87.
16. The composition of claim 13, wherein the genotyping oligonucleotide is a primer-extension oligonucleotide.
17. A method for genotyping the interleukin 13 (IL13) gene of an individual, comprising determining for the two copies of the IL13 gene present in the individual the identity of the nucleotide pair at one or more polymorphic sites (PS) selected from PS1-PS14.
18. The method of claim 17, wherein the determining step comprises:
    (a) isolating from the individual a nucleic acid mixture comprising both copies of the IL13 gene, or a fragment thereof, that are present in the individual;
    (b) amplifying from the nucleic acid mixture a target region containing at least one of the polymorphic sites;
    (c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;
    (d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the polymorphic site; and
    (e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.
19. A method for haplotyping the interleukin 13 (IL13) gene of an individual which comprises determining, for one copy of the IL13 gene present in the individual, the identity of the nucleotide at one or more polymorphic sites (PS) selected from PS1-PS14.
20. The method of claim 19, wherein the determining step comprises
    (a) isolating from the individual a nucleic acid molecule containing only one of the two copies of the IL13 gene, or a fragment thereof, that is present in the individual;
    (b) amplifying from the nucleic acid molecule a target region containing at least one of the polymorphic sites;
    (c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;
    (d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the

reaction, wherein said terminators are complementary to the alternative nucleotides present at the polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

21. A method for predicting a haplotype pair for the interleukin 13 (IL13) gene of an individual comprising:
  (a) identifying an IL13 genotype for the individual at two or more of polymorphic sites selected from PS1-PS14;
  (b) enumerating all possible haplotype pairs which are consistent with the genotype;
  (c) accessing data containing the IL13 haplotype pairs determined in a reference population; and
  (d) assigning a haplotype pair to the individual that is consistent with the data.
22. A method for identifying an association between a trait and at least one genotype or haplotype of the interleukin 13 gene which comprises comparing the frequency of the genotype or haplotype in a population exhibiting the trait with the frequency of the genotype or haplotype in a reference population, wherein the genotype or haplotype comprises a nucleotide pair or nucleotide located at one or more polymorphic sites selected from PS1-PS14, wherein a higher frequency of the genotype or haplotype in the trait population than in the reference population indicates the trait is associated with the genotype or haplotype.
23 The method of claim 22, wherein the haplotype is selected from haplotype numbers 1-21 shown in Table 4.
24. The method of claim 23, wherein the trait is a clinical response to a drug targeting IL13 .
25. A computer system for storing and analyzing polymorphism data for the interleukin 13 gene, comprising:
  (a) a central processing unit (CPU);
  (b) a communication interface;
  (c) a display device;
  (d) an input device; and
  (e) a database containing the polymorphism data;

  wherein the polymorphism data comprises the genotypes and haplotype pairs shown in Table 3 and the haplotypes shown in Table 4.
26. A genome anthology for the interleukin 13 (IL13) gene which comprises IL13 isogenes defined by haplotypes 1-21 shown in Table 4.

POLYMORPHISMS IN THE IL13 GENE

```
GGATCCCCGC TGACAATCTA GAAACAAGCA ACAGGACCCT CTGATGTAGC        50
CATCTGTGCC GCGCCTCTCC GCACCGCCCG CCACGCCTTG GTCCCTGGAG       100
ACCACCCTCC AGGGCAGGGG CTGCCGCTCG GCCGGGCCCG CGGGGTCCCT       150
CGGCCTGACA TGGCCGGTGC TGGAGCGGCA CGTGCGCGCC TCGGCCCCTC       200
GGCCGCTCCC GCCCCTCGCC GGTGCGCACC GGCGCTCGGG GAGCCGCTGG       250
CCCGGGTGTC CAGCCGGCCC TTGCCCTGCC TGGCGCTCGG ACCGCCACCT       300
TTGCCGCCCC CTCGCCAGCC TCCGCAGCTT CCAGACTGGC CGGTCTGCGC       350
GCCCACCCCT GCCTCCCGGA CCGGCCACCG CCGAGGCGCG GAGGAGGGCC       400
CGGCCGCGCA GATCCCGCTT ATCGGCCCAT CTCCCGTTAC ATAAGGCCAC       450
CCCCCTATCT CCGCGGGCCA TCGCCGCCGC AACCGCCGCG CCAGCGCCTT       500
CTCCCACGCG CGGGGGCGCC CCTGCCCACC GCTCCCGGCA GGGCTTTTGG       550
TGGCCATGGG GGATAAGGGG CGTTGACTCA CCCGGGCGGG GCTCCGGGAG       600
TTGCACAGAC CAAGGTAGTT CCCCGCTCCT TCCCCCATCA CGGAGACCCT       650
GTGGGAGATG CCGTGGGCCC TCTACTACAG ATTAGGAAAC AGGCCCGTAG       700
AGGGGTCACA CGGCCAAGTA GCGGCACTCC AGGCACTGGG GGCCCTCGAG       750
GGGAAGGGGC AGACTTCTGG GAGTCAGAGC CAGCAGCTGG GCTGGGAAGC       800
TTCGAGTGTG GACAGAGAGG GTGGGAATGA CGTTCCCTGT GGGAAGAGAG       850
GGTGGCAAGC CTGGGATGCC TCTGAGCGGG AATCCAGCAT GCCTTGTGAG       900
GAGGGTCACA AGCACACCCT TGTGAGGAGG TTGAGCCCCA TCGAGGACAG       950
GACGGAGGGA GCCTGAGCAG GCAGAGAGGG GGCCTGGGGA GGCGCTGGTT      1000
CGGGGAGGAA GTGGGTAGGG GAGAAATCTT GACATCAACA CCCAACAGGC      1050
AAATGCCGTG GCCTCTGCTG TGGGGGTTTC TGGAGGACTT CTAGGAAAAC      1100
                                                   T
GAGGGAAGAG CAGGAAAAGG CGACATGCTG CAGAGACTGG TGAGCAAAGG      1150
GGATCACCCC AAGCCCCAGT GGCACTAGGA ACACTTACAA TCTCTGACCT      1200
GGACTAAGGC TGCCAGCTGG CCCAGTTAAG AGTTTCCCAG AAGGATGGCC      1250
CATACACTTT AAATTAAAGG GGCCAGACAC GTGCACACTA CTTCCAGCCA      1300
CTCTGGAAGC TGAGGTGGGG GGATCGCTTG AGTCTGGGAG TTGGAGGCCA      1350
GCCTAGGCAG GCAACATAGT GAGACCCCAT CTCCAAAAAA ACAAAACAAA      1400
ACAAAACAAA AAAACACCAA AAAAGCTCCC AGAAAGACCT CTGAATCTTT      1450
     G
CTGGATCTCT CAGTGGAGAC CTGGAAATCT GAACTTTGAC AATCCCTCTC      1500
                           G
ACAGTGGGGC CAAGGAGGAA TTAGGCAAGC CAAAAGAAGT GAACTTTACT      1550
CTTCTATTGC CTGTTTGAAT TTTGTATCCA AGCAAGTGTT ACTTAAGTAA      1600
TTTAAGAGAC TGGTTCATCG AAAAAATAAA ACTCCCCAAA TTCCCATAGC      1650
TGGTAGACTG TGGTCACAGC CACAGTGCAC TAAGACTATC TGCTCAGCAC      1700
TTCTGGTGAC CCAAAAGGGT CTGAGGACAG GAGCTCAGAG TTGGGTCAGC      1750
TGTCCAGGTA CTCAGGGTTG TCACAGGCAA AACTGCTGGA ACTCAGGGCA      1800
GCATTGCAAA TGCCACGCCG CTCTCAGGGC CCCTTGCCTG CCGCTGGAAT      1850
                                         T
TAAACCCACC CAGATCTTGG AAACTCTGCC CTGGACCCTT CTCAATAAGT      1900
CCATGAGAAA TCAAACTCTT TCCTTTATGC GACACTGGAT TTTCCACAAA      1950
GTAAAATCAA GATGAGTAAA GATGTGGTTT CTAGATAGTG CCTGAAAAAG      2000
CAGAGACCAT GGTGTCAGGC GTCACCACTT GGGCCTATAA AAGCTGCCAC      2050
AAGACGCCAA GGCCACAAGC CACCCAGCCT ATGCATCCGC TCCTCAATCC      2100
                               C
TCTCCTGTTG GCACTGGGCC TCATGGCGCT TTTGTTGACC ACGGTCATTG      2150
  [Exon 1:2124
CTCTCACTTG CCTTGGCGGC TTTGCCTCCC CAGGCCCTGT GCCTCCCTCT      2200
```

FIGURE 1A

```
ACAGCCCTCA GGGAGCTCAT TGAGGAGCTG GTCAACATCA CCCAGAACCA       2250
                                  A
GAAGGTGAGT GTCGGCTAGC CAGGGTCCTA GCTATGAGGG CTCCAGGGTG       2300
     2255]
GGTGATTCCC AAGATGAGGT CATGAGCAGG CTGGGCCTGG TCCTAAGATG       2350
CCTGTAGGTC AGGAAAAATC TCCATGGACC AAGGCCCGGC CCAGCCATGA       2400
                                       T
GGGAGAGAGG AGCTGGGCTG GGGGGCTCAG CACTGTGGAT GGACCTATGG       2450
AGGTGTCTGG CAGACTCCCC AGGGACTACC TGCTCTCCTG GCCTGGCCTT       2500
GTCTGCCACT GCCAGCTCCT ACTCAGCCAT TCCTGAACAG AGGACAGCAG       2550
AGAAGGGTCC AGCACCCTCC CAGAACCATG TGGCATTTGC CAACTGGATT       2600
TTGACCATAA CAATGCAGCC ATTCTCCCCA GCACCATCAT AGGCCCGCCC       2650
TTACAGGAGG ATTCGTTAGT AGAGTCGCTC CTTGCCCCAC TAGTAACAGC       2700
TCACATGTCT TGAGCACTGC TTACACCAGG CCTGGTGCAC GTGCTTTATG       2750
TGTCATTTCA TCACTGCCAG CCACCTCAAG AGGCAGGTAC GATGAACCCA       2800
TTCTGCTAAG GTTCGTGAGG TTAAGTGACA GAGGCTGGAT TCAAGCCAGG       2850
CCTGGCCAAC ACCAGAGTGT CCATGCTCCT AACTGCAGTG TTCCCTCACC       2900
ATCAGAAGGC AGGGCATTTA ATACACCAGA TCCCCACCGC CTCCCATCTG       2950
ATTTGTCTTG GTCAACATGG CCCAGGCCAC TCCTACTTCA CTCGTCCCCA       3000
CCCTGACCCT TCCCGCAGGC CCCTGTCCTC CTGCCCTGAC TATGGCAAGC       3050
CTTGCATGCA GCTTGTCCCT TACTAGTGGT GTCAATTTTT TTCTCTCAGC       3100
TCCAAGACCC TAAACAGTGG GACCTCACCC CTATGCCTGC TGTTCAAAGC       3150
AGAAAACGAA GCTCAGGAAT GCTGAGGGGC TGCCAGGCCT GCCTCTGTGC       3200
CACACCAGGG ATGCTTGTGG GGCCTGTGCT GGGGCAGACC TGGCCTGGGC       3250
                                                  A
TGCCAGGGCA GGCCCACAAC CCCTGCCAGC ACTCTGCTCA CTGTCACTTT       3300
                             T
GCTCCCACAG CGTCCGCTCT GCAATGGCAG CATGGTATGG AGCATCAACC       3350
      [exon 2:3312..
TGACAGCTGG CATGGTAAGG ACCTTTGGGT GCAGGGAGGA TGGGGCAGAG       3400
               ..3365]
GCTCCAGGCC TTGGGCTTAT CTTCTCTGAG CCTCCCTTCC ATGGCTGGGG       3450
TTCCAAGCAA GCTTCAAGTG CTCTCCTCCC TCCCGCCATA ATCTGGCCCC       3500
TTCCCGCCCA CCACCCAGAC TCACCTGCGC CAGGCATCTC AGCCCCATCT       3550
TCCTGCAGAC TCACAAAAGG CAGCTGCCCA AGCAGGGCCT GACCCCTCGG       3600
TGTCCCCTCC CCACAGTACT GTGCAGCCCT GGAATCCCTG ATCAACGTGT       3650
      [exon 3:3618..
CAGGCTGCAG TGCCATCGAG AAGACCCAGA GGATGCTGAG CGGATTCTGC       3700
CCGCACAAGG TCTCAGCTGG GGTAAGGCAT CCCCCACCCT CTCACACCCA       3750
                   ..3722]
CCCTGCACCC CCTCCTGCCA ACCCTGGGCT CGCTGAAGGG AAGCTGGCTG       3800
AATATCCATG GTGTGTGTCC ACCCAGGGGT GGGGCCATTG TGGCAGCAGG       3850
   T
GACGTGGCCT TCGGGATTTA CAGGATCTGG GCTCAAGGGC TCCTAACTCC       3900
TACCTGGGCC TCAATTTCCA CATCTGTACA GTAGAGGTAC TAACAGTACC       3950
CACCTCATGG GGACTTCCGT GAGGACTGAA TGAGACAGTC CCTGGAAAGC       4000
CCCTGGTTTG TGCGAGTCGT CCCGGCCTCT GGCGTTCTAC TCACGTGCTG       4050
               A                       T          T
ACCTCTTTGT CCTGCAGCAG TTTTCCAGCT TGCATGTCCG AGACACCAAA       4100
    [exon 4:4072..
ATCGAGGTGG CCCAGTTTGT AAAGGACCTG CTCTTACATT TAAAGAAACT       4150
```

FIGURE 1B

```
TTTTCGCGAG GGACGGTTCA ACTGAAACTT CGAAAGCATC ATTATTTGCA        4200
                A
                ..4176]
GAGACAGGAC CTGACTATTG AAGTTGCAGA TTCATTTTTC TTTCTGATGT        4250
CAAAAATGTC TTGGGTAGGC GGGAAGGAGG GTTAGGGAGG GGTAAAATTC        4300
CTTAGCTTAG ACCTCAGCCT GTGCTGCCCG TCTTCAGCCT AGCCGACCTC        4350
AGCCTTCCCC TTGCCCAGGG CTCAGCCTGG TGGGCCTCCT CTGTCCAGGG        4400
CCCTGAGCTC GGTGGACCCA GGGATGACAT GTCCCTACAC CCCTCCCCTG        4450
CCCTAGAGCA CACTGTAGCA TTACAGTGGG TGCCCCCCTT GCCAGACATG        4500
TGGTGGGACA GGGACCCACT TCACACACAG GCAACTGAGG CAGACAGCAG        4550
CTCAGGCACA CTTCTTCTTG GTCTTATTTA TTATTGTGTG TTATTTAAAT        4600
GAGTGTGTTT GTCACCGTTG GGGATTGGGG AAGACTGTGG CTGCTGGCAC        4650
TTGGAGCCAA GGGTTCAGAG ACTCAGGGCC CCAGCACTAA AGCAGTGGAC        4700
CCCAGGAGTC CCTGGTAATA AGTACTGTGT ACAGAATTC                    4739
```

FIGURE 1C

POLYMORPHISMS IN THE CODING SEQUENCE OF IL13

```
ATGGCGCTTT TGTTGACCAC GGTCATTGCT CTCACTTGCC TTGGCGGCTT
TGCCTCCCCA GGCCCTGTGC CTCCCTCTAC AGCCCTCAGG GAGCTCATTG      100
AGGAGCTGGT CAACATCACC CAGAACCAGA AGCGTCCGCT CTGCAATGGC
           A
AGCATGGTAT GGAGCATCAA CCTGACAGCT GGCATGTACT GTGCAGCCCT      200
GGAATCCCTG ATCAACGTGT CAGGCTGCAG TGCCATCGAG AAGACCCAGA
GGATGCTGAG CGGATTCTGC CCGCACAAGG TCTCAGCTGG GTTTTCCAGC      300
TTGCATGTCC GAGACACCAA AATCGAGGTG GCCCAGTTTG TAAAGGACCT
GCTCTTACAT TTAAAGAAAC TTTTTCGCGA GGGACGGTTC AACTGA          396
                                      A
```

FIGURE 2

ISOFORMS OF THE IL13 PROTEIN

```
MALLLTTVIA LTCLGGFASP GPVPPSTALR ELIEELVNIT QNQKRPLCNG
SMVWSINLTA GMYCAALESL INVSGCSAIE KTQRMLSGFC PHKVSAGFSS       100
LHVRDTKIEV AQFVKDLLLH LKKLFREGRF N                           131
                                 Q
```

FIGURE 3

SEQUENCE LISTING

<110> Chew, Anne
Nandabalan, Krishnan
Denton, Richard R.
Stephens, Joel C.

<120> DRUG TARGET ISOGENES: POLYMORPHISMS IN THE INTERLEUKIN 13 GENE

<130> MWH-0011PCT IL13

<140> tba
<141> 2000-09-27

<150> 60/156,489
<151> 1999-09-28

<160> 179

<170> PatentIn Ver. 2.1

<210> 1
<211> 4739
<212> DNA
<213> Homo sapiens

<400> 1
ggatccccgc tgacaatcta gaaacaagca acaggaccct ctgatgtagc catctgtgcc 60
gcgcctctcc gcaccgcccg ccacgccttg gtccctggag accaccctcc agggcagggg 120
ctgccgctcg gccgggcccg cggggtccct cggcctgaca tggccggtgc tggagcggca 180
cgtgcgcgcc tcggcccctc ggccgctccc gcccctcgcc ggtgcgcacc ggcgctcggg 240
gagccgctgg cccgggtgtc cagccggccc ttgccctgcc tggcgctcgg accgccacct 300
ttgccgcccc ctcgccagcc tccgcagctt ccagactggc cggtctgcgc gcccacccct 360
gcctcccgga ccggccaccg ccgaggcgcg gaggagggcc cggccgcgca gatcccgctt 420
atcggcccat ctcccgttac ataaggccac ccccctatct ccgcgggcca tcgccgccgc 480
aaccgccgcg ccagcgcctt ctcccacgcg cgggggcgcc cctgcccacc gctcccggca 540
gggcttttgg tggccatggg ggataagggg cgttgactca cccgggcggg gctccgggag 600
ttgcacagac caaggtagtt ccccgctcct tcccccatca cggagaccct gtgggagatg 660
ccgtgggccc tctactacag attaggaaac aggcccgtag aggggtcaca cggccaagta 720
gcggcactcc aggcactggg ggccctcgag gggaaggggc agacttctgg gagtcagagc 780
cagcagctgg gctgggaagc ttcgagtgtg gacagagagg gtgggaatga cgttccctgt 840
gggaagagag ggtggcaagc ctgggatgcc tctgagcggg aatccagcat gccttgtgag 900
gagggtcaca agcacaccct tgtgaggagg ttgagcccca tcgaggacag gacggaggga 960
gcctgagcag gcagagaggg ggcctgggga ggcgctggtt cggggaggaa gtgggtaggg 1020
gagaaatctt gacatcaaca cccaacaggc aaatgccgtg gcctctgctg tgggggtttc 1080
tggaggactt ctaggaaaac gagggaagag caggaaaagg cgacatgctg cagagactgg 1140
tgagcaaagg ggatcacccc aagccccagt ggcactagga acacttacaa tctctgacct 1200

ggactaaggc tgccagctgg cccagttaag agtttcccag aaggatggcc catacacttt 1260
aaattaaagg ggccagacac gtgcacacta cttccagcca ctctggaagc tgaggtgggg 1320
ggatcgcttg agtctgggag ttggaggcca gcctaggcag gcaacatagt gagaccccat 1380
ctccaaaaaa acaaaacaaa acaaaacaaa aaaacaccaa aaaagctccc agaaagacct 1440
ctgaatcttt ctggatctct cagtggagac ctggaaatct gaactttgac aatccctctc 1500
acagtggggc caaggaggaa ttaggcaagc caaaagaagt gaactttact cttctattgc 1560
ctgtttgaat tttgtatcca agcaagtgtt acttaagtaa tttaagagac tggttcatcg 1620
aaaaaataaa actccccaaa ttcccatagc tggtagactg tggtcacagc cacagtgcac 1680
taagactatc tgctcagcac ttctggtgac ccaaaagggt ctgaggacag gagctcagag 1740
ttgggtcagc tgtccaggta ctcagggttg tcacaggcaa aactgctgga actcagggca 1800
gcattgcaaa tgccacgccg ctctcagggc cccttgcctg ccgctggaat taaacccacc 1860
cagatcttgg aaactctgcc ctggaccctt ctcaataagt ccatgagaaa tcaaactctt 1920
tcctttatgc gacactggat tttccacaaa gtaaaatcaa gatgagtaaa gatgtggttt 1980
ctagatagtg cctgaaaaag cagagaccat ggtgtcaggc gtcaccactt gggcctataa 2040
aagctgccac aagacgccaa ggccacaagc cacccagcct atgcatccgc tcctcaatcc 2100
tctcctgttg gcactgggcc tcatggcgct tttgttgacc acggtcattg ctctcacttg 2160
ccttggcggc tttgcctccc caggccctgt gcctccctct acagccctca gggagctcat 2220
tgaggagctg gtcaacatca cccagaacca gaaggtgagt gtcggctagc cagggtccta 2280
gctatgaggg ctccagggtg ggtgattccc aagatgaggt catgagcagg ctgggcctgg 2340
tcctaagatg cctgtaggtc aggaaaaatc tccatggacc aaggcccggc ccagccatga 2400
gggagagagg agctgggctg gggggctcag cactgtggat ggacctatgg aggtgtctgg 2460
cagactcccc agggactacc tgctctcctg gcctggcctt gtctgccact gccagctcct 2520
actcagccat tcctgaacag aggacagcag agaagggtcc agcaccctcc cagaaccatg 2580
tggcatttgc caactggatt ttgaccataa caatgcagcc attctcccca gcaccatcat 2640
aggcccgccc ttacaggagg attcgttagt agagtcgctc cttgccccac tagtaacagc 2700
tcacatgtct tgagcactgc ttacaccagg cctggtgcac gtgctttatg tgtcatttca 2760
tcactgccag ccacctcaag aggcaggtac gatgaaccca ttctgctaag gttcgtgagg 2820
ttaagtgaca gaggctggat tcaagccagg cctggccaac accagagtgt ccatgctcct 2880
aactgcagtg ttccctcacc atcagaaggc agggcattta atacaccaga tccccaccgc 2940
ctcccatctg atttgtcttg gtcaacatgg cccaggccac tcctacttca ctcgtcccca 3000
ccctgaccct tcccgcaggc ccctgtcctc ctgccctgac tatggcaagc cttgcatgca 3060
gcttgtccct tactagtggt gtcaattttt ttctctcagc tccaagaccc taaacagtgg 3120
gacctcaccc ctatgcctgc tgttcaaagc agaaaacgaa gctcaggaat gctgaggggc 3180
tgccaggcct gcctctgtgc cacaccaggg atgcttgtgg ggcctgtgct ggggcagacc 3240
tggcctgggc tgccagggca ggcccacaac ccctgccagc actctgctca ctgtcacttt 3300
gctcccacag cgtccgctct gcaatggcag catggtatgg agcatcaacc tgacagctgg 3360
catggtaagg acctttgggt gcagggagga tggggcagag gctccaggcc ttgggcttat 3420
cttctctgag cctcccttcc atggctgggg ttccaagcaa gcttcaagtg ctctcctccc 3480
tcccgccata atctggcccc ttcccgccca ccacccagac tcacctgcgc caggcatctc 3540
agccccatct tcctgcagac tcacaaaagg cagctgccca agcagggcct gacccctcgg 3600
tgtcccctcc ccacagtact gtgcagccct ggaatccctg atcaacgtgt caggctgcag 3660
tgccatcgag aagacccaga ggatgctgag cggattctgc ccgcacaagg tctcagctgg 3720
ggtaaggcat cccccaccct ctcacaccca ccctgcaccc cctcctgcca accctgggct 3780
cgctgaaggg aagctggctg aatatccatg gtgtgtgtcc acccaggggt ggggccattg 3840
tggcagcagg gacgtggcct tcgggattta caggatctgg gctcaagggc tcctaactcc 3900
tacctgggcc tcaatttcca catctgtaca gtagaggtac taacagtacc cacctcatgg 3960
ggacttccgt gaggactgaa tgagacagtc cctggaaagc ccctggtttg tgcgagtcgt 4020
cccggcctct ggcgttctac tcacgtgctg acctctttgt cctgcagcag ttttccagct 4080

tgcatgtccg agacaccaaa atcgaggtgg cccagtttgt aaaggacctg ctcttacatt 4140
taaagaaact ttttcgcgag ggacggttca actgaaactt cgaaagcatc attatttgca 4200
gagacaggac ctgactattg aagttgcaga ttcatttttc tttctgatgt caaaaatgtc 4260
ttgggtaggc gggaaggagg gttagggagg ggtaaaattc cttagcttag acctcagcct 4320
gtgctgcccg tcttcagcct agccgacctc agccttcccc ttgcccaggg ctcagcctgg 4380
tgggcctcct ctgtccaggg ccctgagctc ggtggaccca gggatgacat gtccctacac 4440
ccctcccctg ccctagagca cactgtagca ttacagtggg tgccccctt gccagacatg 4500
tggtgggaca gggacccact tcacacacag gcaactgagg cagacagcag ctcaggcaca 4560
cttcttcttg gtcttattta ttattgtgtg ttatttaaat gagtgtgttt gtcaccgttg 4620
gggattgggg aagactgtgg ctgctggcac ttggagccaa gggttcagag actcagggcc 4680
ccagcactaa agcagtggac cccaggagtc cctggtaata agtactgtgt acagaattc 4739

<210> 2
<211> 396
<212> DNA
<213> Homo sapiens

<400> 2
atggcgcttt tgttgaccac ggtcattgct ctcacttgcc ttggcggctt tgcctcccca 60
ggccctgtgc ctccctctac agccctcagg gagctcattg aggagctggt caacatcacc 120
cagaaccaga agcgtccgct ctgcaatggc agcatggtat ggagcatcaa cctgacagct 180
ggcatgtact gtgcagccct ggaatccctg atcaacgtgt caggctgcag tgccatcgag 240
aagacccaga ggatgctgag cggattctgc ccgcacaagg tctcagctgg gttttccagc 300
ttgcatgtcc gagacaccaa aatcgaggtg gcccagtttg taaaggacct gctcttacat 360
ttaaagaaac tttttcgcga gggacggttc aactga 396

<210> 3
<211> 131
<212> PRT
<213> Homo sapiens

<400> 3
Met Ala Leu Leu Leu Thr Thr Val Ile Ala Leu Thr Cys Leu Gly Gly
1 5 10 15

Phe Ala Ser Pro Gly Pro Val Pro Pro Ser Thr Ala Leu Arg Glu Leu
20 25 30

Ile Glu Glu Leu Val Asn Ile Thr Gln Asn Gln Lys Arg Pro Leu Cys
35 40 45

Asn Gly Ser Met Val Trp Ser Ile Asn Leu Thr Ala Gly Met Tyr Cys
50 55 60

Ala Ala Leu Glu Ser Leu Ile Asn Val Ser Gly Cys Ser Ala Ile Glu
65 70 75 80

Lys Thr Gln Arg Met Leu Ser Gly Phe Cys Pro His Lys Val Ser Ala
85 90 95

Gly Phe Ser Ser Leu His Val Arg Asp Thr Lys Ile Glu Val Ala Gln
100 105 110

Phe Val Lys Asp Leu Leu Leu His Leu Lys Lys Leu Phe Arg Glu Gly
115 120 125

Arg Phe Asn
130

<210> 4
<211> 15
<212> DNA
<213> Homo sapiens

<400> 4
aggaaaacga gggaa 15

<210> 5
<211> 15
<212> DNA
<213> Homo sapiens

<400> 5
aggaaaatga gggaa 15

<210> 6
<211> 15
<212> DNA
<213> Homo sapiens

<400> 6
acaaaacaaa aaaac 15

<210> 7
<211> 15
<212> DNA
<213> Homo sapiens

<400> 7
acaaaacgaa aaaac 15

<210> 8
<211> 15
<212> DNA
<213> Homo sapiens

<400> 8
cctggaaatc tgaac 15

<210> 9
<211> 15
<212> DNA
<213> Homo sapiens

<400> 9
cttggagatc tgaac 15

<210> 10
<211> 15
<212> DNA
<213> Homo sapiens

<400> 10
tgcctgccgc tggaa 15

<210> 11
<211> 15
<212> DNA
<213> Homo sapiens

<400> 11
tgcctgctgc tggaa 15

<210> 12
<211> 15
<212> DNA
<213> Homo sapiens

<400> 12
cccagcctat gcatc 15

<210> 13

<211> 15
<212> DNA
<213> Homo sapiens

<400> 13
cccagcccat gcatc 15

<210> 14
<211> 15
<212> DNA
<213> Homo sapiens

<400> 14
agctggtcaa catca 15

<210> 15
<211> 15
<212> DNA
<213> Homo sapiens

<400> 15
agctggtaaa catca 15

<210> 16
<211> 15
<212> DNA
<213> Homo sapiens

<400> 16
accaaggccc ggccc 15

<210> 17
<211> 15
<212> DNA
<213> Homo sapiens

<400> 17
accaaggtcc ggccc 15

<210> 18
<211> 15
<212> DNA
<213> Homo sapiens

<400> 18
tggcctgggc tgcca 15

<210> 19
<211> 15
<212> DNA
<213> Homo sapiens

<400> 19
tggcctgagc tgcca 15

<210> 20
<211> 15
<212> DNA
<213> Homo sapiens

<400> 20
ctgccagcac tctgc 15

<210> 21
<211> 15
<212> DNA
<213> Homo sapiens

<400> 21
ctgccagtac tctgc 15

<210> 22
<211> 15
<212> DNA
<213> Homo sapiens

<400> 22
gaatatccat ggtgt 15

<210> 23
<211> 15
<212> DNA
<213> Homo sapiens

<400> 23
gaatatctat ggtgt 15

<210> 24
<211> 15
<212> DNA
<213> Homo sapiens

<400> 24
tttgtgcgag tcgtc 15

<210> 25
<211> 15
<212> DNA
<213> Homo sapiens

<400> 25
tttgtgcaag tcgtc 15

<210> 26
<211> 15
<212> DNA
<213> Homo sapiens

<400> 26
cctctggcgt tctac 15

<210> 27
<211> 15
<212> DNA
<213> Homo sapiens

<400> 27
cctctggtgt tctac 15

<210> 28
<211> 15
<212> DNA
<213> Homo sapiens

<400> 28
ctactcacgt gctga 15

<210> 29

<211> 15
<212> DNA
<213> Homo sapiens

<400> 29
ctactcatgt gctga 15

<210> 30
<211> 15
<212> DNA
<213> Homo sapiens

<400> 30
gagggacggt tcaac 15

<210> 31
<211> 15
<212> DNA
<213> Homo sapiens

<400> 31
gagggacagt tcaac 15

<210> 32
<211> 15
<212> DNA
<213> Homo sapiens

<400> 32
acttctagga aaacg 15

<210> 33
<211> 15
<212> DNA
<213> Homo sapiens

<400> 33
ctgctcttcc ctcgt 15

<210> 34
<211> 15
<212> DNA
<213> Homo sapiens

<400> 34
acttctagga aaatg 15

<210> 35
<211> 15
<212> DNA
<213> Homo sapiens

<400> 35
ctgctcttcc ctcat 15

<210> 36
<211> 15
<212> DNA
<213> Homo sapiens

<400> 36
aacaaaacaa aacaa 15

<210> 37
<211> 15
<212> DNA
<213> Homo sapiens

<400> 37
tttggtgttt ttttg 15

<210> 38
<211> 15
<212> DNA
<213> Homo sapiens

<400> 38
aacaaaacaa aacga 15

<210> 39
<211> 15
<212> DNA
<213> Homo sapiens

<400> 39
tttggtgttt tttcg 15

<210> 40
<211> 15
<212> DNA
<213> Homo sapiens

<400> 40
tggagacctg gaaat 15

<210> 41
<211> 15
<212> DNA
<213> Homo sapiens

<400> 41
gtcaaagttc agatt 15

<210> 42
<211> 15
<212> DNA
<213> Homo sapiens

<400> 42
tggagacctg gaagt 15

<210> 43
<211> 15
<212> DNA
<213> Homo sapiens

<400> 43
gtcaaagttc agact 15

<210> 44
<211> 15
<212> DNA
<213> Homo sapiens

<400> 44
gcccсttgcc tgccg 15

<210> 45

<211> 15
<212> DNA
<213> Homo sapiens

<400> 45
gtttaattcc agcgg 15

<210> 46
<211> 15
<212> DNA
<213> Homo sapiens

<400> 46
gcccttgcc tgctg 15

<210> 47
<211> 15
<212> DNA
<213> Homo sapiens

<400> 47
gtttaattcc agcag 15

<210> 48
<211> 15
<212> DNA
<213> Homo sapiens

<400> 48
aagccaccca gccta 15

<210> 49
<211> 15
<212> DNA
<213> Homo sapiens

<400> 49
ggagcggatg catag 15

<210> 50
<211> 15
<212> DNA
<213> Homo sapiens

<400> 50
aagccaccca gccca 15

<210> 51
<211> 15
<212> DNA
<213> Homo sapiens

<400> 51
ggagcggatg catgg 15

<210> 52
<211> 15
<212> DNA
<213> Homo sapiens

<400> 52
ttgaggagct ggtca 15

<210> 53
<211> 15
<212> DNA
<213> Homo sapiens

<400> 53
tctgggtgat gttga 15

<210> 54
<211> 15
<212> DNA
<213> Homo sapiens

<400> 54
ttgaggagct ggtaa 15

<210> 55
<211> 15
<212> DNA
<213> Homo sapiens

<400> 55
tctgggtgat gttta 15

<210> 56
<211> 15
<212> DNA
<213> Homo sapiens

<400> 56
ccatggacca aggcc 15

<210> 57
<211> 15
<212> DNA
<213> Homo sapiens

<400> 57
atggctgggc cgggc 15

<210> 58
<211> 15
<212> DNA
<213> Homo sapiens

<400> 58
ccatggacca aggtc 15

<210> 59
<211> 15
<212> DNA
<213> Homo sapiens

<400> 59
atggctgggc cggac 15

<210> 60
<211> 15
<212> DNA
<213> Homo sapiens

<400> 60
cagacctggc ctggg 15

<210> 61

<211> 15
<212> DNA
<213> Homo sapiens

<400> 61
ctgccctggc agccc 15

<210> 62
<211> 15
<212> DNA
<213> Homo sapiens

<400> 62
cagacctggc ctgag 15

<210> 63
<211> 15
<212> DNA
<213> Homo sapiens

<400> 63
ctgccctggc agctc 15

<210> 64
<211> 15
<212> DNA
<213> Homo sapiens

<400> 64
caacccctgc cagca 15

<210> 65
<211> 15
<212> DNA
<213> Homo sapiens

<400> 65
cagtgagcag agtgc 15

<210> 66
<211> 15
<212> DNA
<213> Homo sapiens

<400> 66
caacccctgc cagta 15

<210> 67
<211> 15
<212> DNA
<213> Homo sapiens

<400> 67
cagtgagcag agtac 15

<210> 68
<211> 15
<212> DNA
<213> Homo sapiens

<400> 68
ctggctgaat atcca 15

<210> 69
<211> 15
<212> DNA
<213> Homo sapiens

<400> 69
ggacacacac catgg 15

<210> 70
<211> 15
<212> DNA
<213> Homo sapiens

<400> 70
ctggctgaat atcta 15

<210> 71
<211> 15
<212> DNA
<213> Homo sapiens

<400> 71
ggacacacac catag 15

<210> 72
<211> 15
<212> DNA
<213> Homo sapiens

<400> 72
ccctggtttg tgcga 15

<210> 73
<211> 15
<212> DNA
<213> Homo sapiens

<400> 73
ggccgggacg actcg 15

<210> 74
<211> 15
<212> DNA
<213> Homo sapiens

<400> 74
ccctggtttg tgcaa 15

<210> 75
<211> 15
<212> DNA
<213> Homo sapiens

<400> 75
ggccgggacg acttg 15

<210> 76
<211> 15
<212> DNA
<213> Homo sapiens

<400> 76
tcccggcctc tggcg 15

<210> 77

<211> 15
<212> DNA
<213> Homo sapiens

<400> 77
acgtgagtag aacgc 15

<210> 78
<211> 15
<212> DNA
<213> Homo sapiens

<400> 78
tcccggcctc tggtg 15

<210> 79
<211> 15
<212> DNA
<213> Homo sapiens

<400> 79
acgtgagtag aacac 15

<210> 80
<211> 15
<212> DNA
<213> Homo sapiens

<400> 80
ggcgttctac tcacg 15

<210> 81
<211> 15
<212> DNA
<213> Homo sapiens

<400> 81
aagaggtcag cacgt 15

<210> 82
<211> 15
<212> DNA
<213> Homo sapiens

<400> 82
ggcgttctac tcatg 15

<210> 83
<211> 15
<212> DNA
<213> Homo sapiens

<400> 83
aagaggtcag cacat 15

<210> 84
<211> 15
<212> DNA
<213> Homo sapiens

<400> 84
tttcgcgagg gacgg 15

<210> 85
<211> 15
<212> DNA
<213> Homo sapiens

<400> 85
gtttcagttg aaccg 15

<210> 86
<211> 15
<212> DNA
<213> Homo sapiens

<400> 86
tttcgcgagg gacag 15

<210> 87
<211> 15
<212> DNA
<213> Homo sapiens

<400> 87
gtttcagttg aactg 15

<210> 88
<211> 10
<212> DNA
<213> Homo sapiens

<400> 88
tctaggaaaa 10

<210> 89
<211> 10
<212> DNA
<213> Homo sapiens

<400> 89
ctcttccctc 10

<210> 90
<211> 10
<212> DNA
<213> Homo sapiens

<400> 90
aaaacaaaac 10

<210> 91
<211> 10
<212> DNA
<213> Homo sapiens

<400> 91
ggtgtttttt 10

<210> 92
<211> 9
<212> DNA
<213> Homo sapiens

<400> 92
agacctgga 9

<210> 93

<211> 10
<212> DNA
<213> Homo sapiens

<400> 93
aaagttcaga 10

<210> 94
<211> 10
<212> DNA
<213> Homo sapiens

<400> 94
ccttgcctgc 10

<210> 95
<211> 10
<212> DNA
<213> Homo sapiens

<400> 95
taattccagc 10

<210> 96
<211> 10
<212> DNA
<213> Homo sapiens

<400> 96
ccacccagcc 10

<210> 97
<211> 10
<212> DNA
<213> Homo sapiens

<400> 97
gcggatgcat 10

<210> 98
<211> 10
<212> DNA
<213> Homo sapiens

<400> 98
aggagctggt 10

<210> 99
<211> 10
<212> DNA
<213> Homo sapiens

<400> 99
gggtgatgtt 10

<210> 100
<211> 10
<212> DNA
<213> Homo sapiens

<400> 100
tggaccaagg 10

<210> 101
<211> 10
<212> DNA
<213> Homo sapiens

<400> 101
gctgggccgg 10

<210> 102
<211> 10
<212> DNA
<213> Homo sapiens

<400> 102
acctggcctg 10

<210> 103
<211> 10
<212> DNA
<213> Homo sapiens

<400> 103
ccctggcagc 10

<210> 104
<211> 10
<212> DNA
<213> Homo sapiens

<400> 104
cccctgccag 10

<210> 105
<211> 10
<212> DNA
<213> Homo sapiens

<400> 105
tgagcagagt 10

<210> 106
<211> 10
<212> DNA
<213> Homo sapiens

<400> 106
gctgaatatc 10

<210> 107
<211> 10
<212> DNA
<213> Homo sapiens

<400> 107
cacacaccat 10

<210> 108
<211> 10
<212> DNA
<213> Homo sapiens

<400> 108
tggtttgtgc 10

<210> 109

<211> 10
<212> DNA
<213> Homo sapiens

<400> 109
cgggacgact 10

<210> 110
<211> 10
<212> DNA
<213> Homo sapiens

<400> 110
cggcctctgg 10

<210> 111
<211> 10
<212> DNA
<213> Homo sapiens

<400> 111
tgagtagaac 10

<210> 112
<211> 10
<212> DNA
<213> Homo sapiens

<400> 112
gttctactca 10

<210> 113
<211> 10
<212> DNA
<213> Homo sapiens

<400> 113
aggtcagcac 10

<210> 114
<211> 10
<212> DNA
<213> Homo sapiens

<400> 114
cgcgagggac 10

<210> 115
<211> 10
<212> DNA
<213> Homo sapiens

<400> 115
tcagttgaac 10

<210> 116
<211> 20
<212> DNA
<213> Homo sapiens

<400> 116
agcctgagca ggcagagagg 20

<210> 117
<211> 20
<212> DNA
<213> Homo sapiens

<400> 117
cttggcccca ctgtgagagg 20

<210> 118
<211> 23
<212> DNA
<213> Homo sapiens

<400> 118
ccagacacgt gcacactact tcc 23

<210> 119
<211> 22
<212> DNA
<213> Homo sapiens

<400> 119
caaccctgag tacctggaca gc 22

<210> 120
<211> 22
<212> DNA
<213> Homo sapiens

<400> 120
gccaaggagg aattaggcaa gc 22

<210> 121
<211> 22
<212> DNA
<213> Homo sapiens

<400> 121
agctccctga gggctgtaga gg 22

<210> 122
<211> 24
<212> DNA
<213> Homo sapiens

<400> 122
ccctggaccc ttctcaataa gtcc 24

<210> 123
<211> 21
<212> DNA
<213> Homo sapiens

<400> 123
cagctcctct ctccctcatg g 21

<210> 124
<211> 24
<212> DNA
<213> Homo sapiens

<400> 124
cctttatgcg acactggatt ttcc 24

<210> 125

<211> 22
<212> DNA
<213> Homo sapiens

<400> 125
aggccaggag agcaggtagt cc 22

<210> 126
<211> 23
<212> DNA
<213> Homo sapiens

<400> 126
caaagcagaa aacgaagctc agg 23

<210> 127
<211> 23
<212> DNA
<213> Homo sapiens

<400> 127
tttgtgagtc tgcaggaaga tgg 23

<210> 128
<211> 22
<212> DNA
<213> Homo sapiens

<400> 128
gttccaagca agcttcaagt gc 22

<210> 129
<211> 22
<212> DNA
<213> Homo sapiens

<400> 129
tgaggcccag gtaggagtta gg 22

<210> 130
<211> 21
<212> DNA
<213> Homo sapiens

<400> 130
cctacctggg cctcaatttc c 21

<210> 131
<211> 21
<212> DNA
<213> Homo sapiens

<400> 131
gtgtgaagtg ggtccctgtc c 21

<210> 132
<211> 20
<212> DNA
<213> Homo sapiens

<400> 132
gcctgagcag gcagagaggg 20

<210> 133
<211> 20
<212> DNA
<213> Homo sapiens

<400> 133
ttggccccac tgtgagaggg 20

<210> 134
<211> 22
<212> DNA
<213> Homo sapiens

<400> 134
agacacgtgc acactacttc ca 22

<210> 135
<211> 21
<212> DNA
<213> Homo sapiens

<400> 135
accctgagta cctggacagc t 21

<210> 136
<211> 21
<212> DNA
<213> Homo sapiens

<400> 136
ccccaaattc ccatagctgg t 21

<210> 137
<211> 20
<212> DNA
<213> Homo sapiens

<400> 137
gcaatgaccg tggtcaacaa 20

<210> 138
<211> 20
<212> DNA
<213> Homo sapiens

<400> 138
gcgacactgg attttccaca 20

<210> 139
<211> 20
<212> DNA
<213> Homo sapiens

<400> 139
ggccttggtc catggagatt 20

<210> 140
<211> 22
<212> DNA
<213> Homo sapiens

<400> 140
gtgcctgaaa aagcagagac ca 22

<210> 141

<211> 20
<212> DNA
<213> Homo sapiens

<400> 141
aggtccatcc acagtgctga 20

<210> 142
<211> 19
<212> DNA
<213> Homo sapiens

<400> 142
ctgcctctgt gccacacca 19

<210> 143
<211> 19
<212> DNA
<213> Homo sapiens

<400> 143
attatggcgg gagggagga 19

<210> 144
<211> 20
<212> DNA
<213> Homo sapiens

<400> 144
ctccctcccg ccataatctg 20

<210> 145
<211> 20
<212> DNA
<213> Homo sapiens

<400> 145
taaatcccga aggccacgtc 20

<210> 146
<211> 20
<212> DNA
<213> Homo sapiens

<400> 146
acccacctca tggggacttc 20

<210> 147
<211> 20
<212> DNA
<213> Homo sapiens

<400> 147
ggtcggctag gctgaagacg 20

<210> 148
<211> 23
<212> DNA
<213> Homo sapiens

<400> 148
gaaacaagca acaggaccct ctg 23

<210> 149
<211> 25
<212> DNA
<213> Homo sapiens

<400> 149
ggaactacct tggtctgtgc aactc 25

<210> 150
<211> 22
<212> DNA
<213> Homo sapiens

<400> 150
gggataaggg gcgttgactc ac 22

<210> 151
<211> 26
<212> DNA
<213> Homo sapiens

<400> 151
ccatccttct gggaaactct taactg 26

<210> 152
<211> 26
<212> DNA
<213> Homo sapiens

<400> 152
gtttctggag gacttctagg aaaacg 26

<210> 153
<211> 24
<212> DNA
<213> Homo sapiens

<400> 153
tgaccacagt ctaccagcta tggg 24

<210> 154
<211> 24
<212> DNA
<213> Homo sapiens

<400> 154
gcaacatagt gagaccccat ctcc 24

<210> 155
<211> 24
<212> DNA
<213> Homo sapiens

<400> 155
gcgtcttgtg gcagctttta tagg 24

<210> 156
<211> 23
<212> DNA
<213> Homo sapiens

<400> 156
ccccaaattc ccatagctgg tag 23

<210> 157

<211> 22
<212> DNA
<213> Homo sapiens

<400> 157
gccgacactc tccttctggt tc 22

<210> 158
<211> 25
<212> DNA
<213> Homo sapiens

<400> 158
gatagtgcct gaaaaagcag agacc 25

<210> 159
<211> 23
<212> DNA
<213> Homo sapiens

<400> 159
cctccatagg tccatccaca gtg 23

<210> 160
<211> 21
<212> DNA
<213> Homo sapiens

<400> 160
tcacccctat gcctgctgtt c 21

<210> 161
<211> 24
<212> DNA
<213> Homo sapiens

<400> 161
aggtaggagt taggagccct tgag 24

<210> 162
<211> 21
<212> DNA
<213> Homo sapiens

<400> 162
tgagacagtc cctggaaagc c 21

<210> 163
<211> 24
<212> DNA
<213> Homo sapiens

<400> 163
aaggaatttt acccctccct aacc 24

<210> 164
<211> 24
<212> DNA
<213> Homo sapiens

<400> 164
aagcaacagg accctctgat gtag 24

<210> 165
<211> 24
<212> DNA
<213> Homo sapiens

<400> 165
aactaccttg gtctgtgcaa ctcc 24

<210> 166
<211> 23
<212> DNA
<213> Homo sapiens

<400> 166
ttctaggaaa acgagggaag agc 23

<210> 167
<211> 24
<212> DNA
<213> Homo sapiens

<400> 167
tgaccacagt ctaccagcta tggg 24

<210> 168
<211> 26
<212> DNA
<213> Homo sapiens

<400> 168
gaatctttct ggatctctca gtggag 26

<210> 169
<211> 21
<212> DNA
<213> Homo sapiens

<400> 169
caagtggtga cgcctgactc c 21

<210> 170
<211> 25
<212> DNA
<213> Homo sapiens

<400> 170
aagactatct gctcagcact tctgg 25

<210> 171
<211> 23
<212> DNA
<213> Homo sapiens

<400> 171
tgatgttgac cagctcctca atg 23

<210> 172
<211> 20
<212> DNA
<213> Homo sapiens

<400> 172
acaagccacc cagcctatgc 20

<210> 173

<211> 30
<212> DNA
<213> Homo sapiens

<400> 173
catccacagt gctgagcccc ccagcccagc 30

<210> 174
<211> 20
<212> DNA
<213> Homo sapiens

<400> 174
cacaccaggg atgcttgtgg 20

<210> 175
<211> 21
<212> DNA
<213> Homo sapiens

<400> 175
ctggaagatg ggctgagat g 21

<210> 176
<211> 33
<212> DNA
<213> Homo sapiens

<400> 176
tccctgatca acgtgtcagg ctgcagtgcc atc 33

<210> 177
<211> 22
<212> DNA
<213> Homo sapiens

<400> 177
tattcagcca gcttcccttc ag 22

<210> 178
<211> 20
<212> DNA
<213> Homo sapiens

<400> 178
cctggtttgt gcgagtcgtc 20

<210> 179
<211> 20
<212> DNA
<213> Homo sapiens

<400> 179
cccctcccta accctccttc 20